

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 23-08-2004	2. REPORT TYPE Conference Proceedings	3. DATES COVERED (From – To) 14 June 2004 - 19 June 2004
--	---	--

4. TITLE AND SUBTITLE Complex Systems: Control and Modeling Problems	5a. CONTRACT NUMBER FA8655-03-1-5088
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S) Conference Committee	5d. PROJECT NUMBER
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Russian Academy of Sciences 61, Sadovaya str. Samara 443020 Russia	8. PERFORMING ORGANIZATION REPORT NUMBER N/A
--	--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 BOX 14 FPO 09499-0014	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) CSP 03-5088

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited. (approval given by local Public Affairs Office)

13. SUPPLEMENTARY NOTES
Copyright 2004. The Department of Defense has permission to use for government purposes only. All other rights are reserved by the copyright holder.

14. ABSTRACT

The Final Proceedings for Complex Systems: Control and Modeling Problems, 14 June 2004 - 19 June 2004

This is a computer science conference broadly covering topics related to modeling and control of complex systems and 'systems of systems'. Specific topics to be presented include: Open systems: Control and Modeling; Complex Systems: Information Interaction Models; Information Assurance in Complex Systems; System Analysis and Control Theory; Ontology analysis and synthesis; Multi-Agent Systems; Complex Engineering Systems and Enterprises management; Emergency Control; Control and Measurement in Complex Technical Systems; New Information Technologies.

15. SUBJECT TERMS
EOARD, Modeling & Simulation, Computer Science, Multi Agent Systems, Complex Systems, Open Systems

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UL	18, NUMBER OF PAGES 483	19a. NAME OF RESPONSIBLE PERSON /Signed/PAUL LOSIEWICZ, Ph. D.
a. REPORT UNCLAS	b. ABSTRACT UNCLAS	c. THIS PAGE UNCLAS			19b. TELEPHONE NUMBER (Include area code) +44 20 7514 4474



Российская Академия Наук

**ПРОБЛЕМЫ УПРАВЛЕНИЯ
И МОДЕЛИРОВАНИЯ
В СЛОЖНЫХ СИСТЕМАХ**

**COMPLEX SYSTEMS:
CONTROL AND
MODELING PROBLEMS**

Труды
VI Международная конференция
14-17 июня 2004
Самара
Россия

Proceedings
The VI International Conference
June 14-17, 2004
Samara
Russia

Российская академия наук
Самарский научный центр
Институт проблем управления сложными системами

Труды VI Международной конференции
Proceedings of the VI International Conference

**ПРОБЛЕМЫ УПРАВЛЕНИЯ
И МОДЕЛИРОВАНИЯ В СЛОЖНЫХ СИСТЕМАХ**

**COMPLEX SYSTEMS:
CONTROL AND MODELING PROBLEMS**

14-17 июня 2004, Самара, Россия

June 14-17, 2004, Samara, Russia

Редакторы:

академик Е.А. Федосов

академик Н.А. Кузнецов

профессор В.А. Виттих

Editors:

academician E.A. Fedosov

academician N.A. Kuznetsov

professor V.A. Vittikh

Проблемы управления и моделирования в сложных системах: Труды VI Международной конференции / Под ред.: **акад. Е.А. Федосова**, **акад. Н.А. Кузнецова**, проф. **В.А. Виттиха**. - Самара: Самарский научный центр РАН, 2004. - 480 с.

ISBN 5-93424-140-0

В предлагаемом сборнике содержатся труды VI Международной конференции «Проблемы управления и моделирования в сложных системах», распределенные по следующим направлениям: системный анализ и теория управления, модели информационного взаимодействия в сложных системах, новые информационные технологии, мультиагентные системы, региональное управление, управление и измерения в сложных технических системах.

ISBN 5-93424-140-0

- © Институт проблем управления сложными системами РАН, составление, оформление, 2004
- © Все права принадлежат авторам публикуемых работ, 2004

The VI International Conference “Complex Systems: Control and Modeling Problems”, was held June 14-17, 2004 at Samara (Russia) by the International Association for Mathematics and Computers in Simulation (IMACS), National Committee of Automatic Control of Russia, Samara Scientific Center of RAS, Institute for the Control of Complex Systems (ICCS) of RAS at support of Administration of the Samara Region and at participation Volga Region Academy of Telecommunications and Informatic.

The Conference was funded to the project № T4254 of Federal Special Program “State Support of Integration of High Education and Fundamental Science on 2002-2006”.

The sponsors of the Conference were MagentA Corporation Ltd. (London, UK), JSC AutoVAZ, Togliattikauchuk LLC, JSC Kuibyshevazot, JSC Togliattiazot and MEC Knowledge Genesis.

Organizers wish to thank the following for their contribution to the success of this conference: European Office of Aerospace Research and Development (London, UK), Air Force Office of Scientific Research (London, UK), United States Air Force Research Laboratory (London, UK).

Presented on this conference reports consider actual problems in the fields of:

- System Analysis and Control Theory;
- Complex Systems: Information Interaction Models;
- New Information Technologies;
- Multi-Agent Systems;
- Regional Management;
- Control and Measurement in Complex Technical Systems.

The conference was participated by the scientists from universities and research institutes of Great Britain, USA, Germany and Belgium. The Russian participants represented Russian academy of sciences, universities, research and development centers and enterprises.

VI Международная конференция «Проблемы управления и моделирования в сложных системах» проводилась в г. Самаре (Россия) 14-17 июня 2004 года Международной ассоциацией по математическому и компьютерному моделированию (IMACS), Российским Национальным комитетом по автоматическому управлению, Самарским научным центром РАН, Институтом проблем управления сложными системами (ИПУСС) РАН при поддержке Администрации Самарской области и при участии Поволжской государственной академии телекоммуникаций и информатики.

Конференция получила финансовую поддержку Федеральной целевой программы «Государственная поддержка интеграции высшего образования и фундаментальной науки на 2002-2006 годы» по проекту № Т4254.

В качестве спонсоров Конференции выступили корпорация MagentA Ltd. (г. Лондон, Великобритания), АО «АвтоВАЗ», ООО «Тольяттикаучук», ЗАО «Куйбышевазот», ОАО «Тольяттиазот» и ООО НПК «Генезис знаний».

Организаторы выражают благодарность за вклад в успешное проведение конференции следующим организациям: Европейскому отделу аэрокосмических исследований и разработок (г. Лондон, Великобритания), Отделу научных исследований Военно-воздушных сил (г. Лондон, Великобритания), Исследовательской лаборатории Военно-воздушных сил США (г. Лондон, Великобритания).

На Конференции были представлены доклады по следующим секциям:

- системный анализ и теория управления;
- модели информационного взаимодействия в сложных системах;
- новые информационные технологии;
- мультиагентные системы;
- региональное управление;
- управление и измерения в сложных технических системах.

В Конференции приняли участие ученые из университетов и научных учреждений Великобритании, США, Германии и Бельгии. Отечественные ученые представляли Российскую академию наук, вузы, научно-производственные объединения и промышленные предприятия страны.

ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

1. *Федосов Е.А.* - академик РАН, председатель Программного комитета
2. *Виттих В.А.* - д.т.н., заместитель председателя, ИПУСС РАН, г. Самара
3. *Боровик С.Ю.* - к.т.н., ученый секретарь, ИПУСС РАН, г. Самара
4. *Аншаков Г.П.* - член-корреспондент РАН
5. *Брандт Д.* - доктор наук, Университет в г. Аахен, Германия
6. *Васильев С.Н.* - член-корреспондент РАН
7. *Вишневецкий Р.* - президент IMACS, профессор Университета Радгерса, г. Нью-Йорк, США
8. *Зибарев А.Г.* - член-корреспондент РАН
9. *Каданников В.В.* - председатель Совета директоров АО «АВТОВАЗ»
10. *Кузнецов Н.А.* - академик РАН
11. *Лейфер Л.* - профессор Стенфордского университета, г. Пало Альто, США
12. *Нариньяни А.С.* - к.ф.-м.н., Российский НИИ искусственного интеллекта, г. Москва
13. *Прангивили И.В.* - академик АН Грузии
14. *Ржевский Г.А.* - профессор Университета Brunel, г. Лондон, Великобритания
15. *Себряков Г.Г.* - член-корреспондент РАН
16. *Теряев Е.Д.* - член-корреспондент РАН
17. *Титов К.А.* - д.э.н., губернатор Самарской области
18. *Тьугу Э.Х.* - профессор Королевского технического университета, г. Стокгольм, Швеция
19. *Хасебе М.* - профессор университета Уценомия, префектура Точиги, Япония
20. *Шорин В.П.* - академик РАН

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

1. *Виттих В.А.* - д.т.н., председатель Организационного комитета, ИПУСС РАН, г. Самара
2. *Смирнов С.В.* - д.т.н., заместитель председателя, ИПУСС РАН, г. Самара
3. *Боровик С.Ю.* - к.т.н., ученый секретарь, ИПУСС РАН, г. Самара
4. *Андреев В.А.* - д.т.н., Поволжская государственная академия телекоммуникаций и информатики (г. Самара)
5. *Диллигенский Н.В.* - д.т.н., Самарский государственный технический университет
6. *Жеребин А.М.* - д.т.н., ГосНИИ авиационных систем, г. Москва
7. *Золотухин Ю.Н.* - д.т.н., Институт автоматике и электрометрии СО РАН, г. Новосибирск
8. *Ильясов Б.Г.* - д.т.н., Уфимский государственный авиационный технический университет
9. *Карпов В.М.* - АО «АВТОВАЗ», г. Тольятти
10. *Клещев А.С.* - д.ф.-м.н., Институт автоматике и процессов управления ДВО РАН, г. Владивосток
11. *Кузнецов С.В.* - ИПУСС РАН, г. Самара
12. *Любецкий В.А.* - д.ф.-м.н., Институт проблем передачи информации РАН, г. Москва
13. *Мерджанов В.Р.* - к.х.н., Министерство промышленности и технологий Самарской области
14. *Мосыченко Ю.И.* - к.т.н., заместитель руководителя аппарата Администрации Самарской области
15. *Санчугов В.И.* - д.т.н., Самарский научный центр РАН
16. *Скобелев П.О.* - д.т.н., НПК «Генезис знаний», г. Самара
17. *Шубин А.Н.* - к.т.н., Институт проблем управления РАН, г. Москва

CONTENTS СОДЕРЖАНИЕ

Plenary Papers

Пленарные доклады

1. *G. Rzevski* 1
Designing complex engineering systems
2. *Н.А. Кузнецов, М.С. Гельфанд, В.А. Любецкий* 5
Биоинформатика бактерий – основные задачи и характер моделей
3. *N.H. Silbert* 12
A consideration of public choice theory for the analysis of change in public policy using multi-agent systems
4. *А.М. Жеребин, В.В. Кронова, В.Г. Баров* 18
Критериальные аспекты принятия решений в задачах проектирования, оценки эффективности, оптимизации состава сложных технических систем
5. *A.V. Smirnov, M.P. Pashkin, N.V. Chilov, T.V. Levashova, A.A. Krizhanovsky* 27
Context-based knowledge logistics for disaster management
6. *V.A. Buntux* 37
Знания, основанные на понимании, в процессах принятия решений
7. *V.A. Skormin, D.H. Summerville, O.I. Shiryayeva, J.S. Moronski* 45
Analysis of active response in the immune system with computer network considerations

System Analysis and Control Theory

Системный анализ и теория управления

1. *Б.П. Топоров, В.М. Горлов* 57
Методы решения задачи типажа сложных технических систем
2. *Э.Я. Панопорт* 64
Методы структурной теории в задачах синтеза систем управления
3. *R.H. Rigby, V.A. Zarubin* 76
Effective management techniques for international engineering projects
4. *Н.В. Дилигенский, А.П. Ефимов* 84
Системный анализ и модели информационного воздействия механизмов эволюционного развития
5. *Н.В. Дилигенский, М.В. Цапенко* 96
Математическое моделирование и обобщенное оценивание эффективности производственно-экономических систем
6. *М.Б. Гузаиров, Р.А. Бадамин, Б.Г. Ильясов, И.Б. Герасимова, Н.В. Хасанова* 107
Синергетический подход к исследованию организационных систем
7. *В.Б. Тарасов* 113
Самообучающиеся предприятия

8.	<i>В.В. Емельянов, П.В. Афонин</i> Модели искусственной жизни для решения оптимизационных задач в сложных дискретных системах	123
9.	<i>В.В. Емельянов, Ю.О. Шепитько</i> Модели в системах адаптивного контроля знаний	130
10.	<i>В.К. Захаров, А.Д. Яшин</i> Динамическая модель самоорганизующегося нома	137
11.	<i>С.В. Архангельский, Н.А. Симон</i> Задача синтеза структуры логистического центра	145

Complex Systems: Information Interaction Models

Модели информационного взаимодействия в сложных системах

1.	<i>К.Ю. Горбунов, В.А. Любецкий</i> Модель регуляции биосинтеза	151
2.	<i>В.А. Любецкий, А.В. Селиверстов</i> Геометрический метод поиска клики в графе и его применение для выделения сигнала	154
3.	<i>Е.В. Любецкая, Л.А. Леонтьев, М.А. Ширишин, В.А. Любецкий</i> Алгоритм поиска сигнала со сложной структурой	158
4.	<i>А.А. Миронов</i> Биоинформатика: от обслуживания эксперимента к получению биологически значимых результатов	163
5.	<i>С.Н. Истомина</i> Информационное управление процессом изучения сложной реакции	166

New information technologies

Новые информационные технологии

1.	<i>О.А. Невзорова, В.Н. Невзоров</i> Инструментальная система визуального проектирования онтологий «OntoEditor»	170
2.	<i>И.Д. Гофман, Д.А. Инишев, А.В. Шурбаков, Л.Г. Романов</i> Моделирование ресурсных ограничений при структурном недоопределенном планировании в технологии Time-EX	176
3.	<i>Ю.В. Костов, Д.А. Липовой, П.Г. Мамонтов, Е.С. Петров</i> Новая версия универсального решателя UniCalc: возможности и перспективы развития	183
4.	<i>А. Артиба, А. Донтэн, С.И. Ясиновский</i> ALIX: метод графического моделирования процессов в сложных дискретных системах	190

5. *М.Б. Гузаиров, Р.А. Бадамишин, Б.Г. Ильясов, Л.Р. Черняховская, Р.А. Шкундина* 197
Интеграция моделей знаний на основе объектно-когнитивного анализа
6. *Н.И. Юсупова, Д.В. Попов, И.А. Вайнерман, Б. Кёних-Риз* 203
RDF-технологии предоставления сервисов
мобильным пользователям в беспроводных сетях
7. *А.М. Кистанов* 209
Онтологический анализ и модели информационного взаимодействия
платежных систем
8. *О.И. Боровикова, С.В. Булгаков, Ю.А. Загорулько, Е.В. Сидорова,
Ю.П. Холюшкин* 215
Концепция интеллектуального Интернет-портала знаний
для доступа к информационным ресурсам по археологии и этнографии
9. *Г.Б. Загорулько* 221
Представление знаний в системе сопровождения курса лечения элементозов
10. *В.В. Тарасевич* 227
Использование метода недоопределенных вычислений
для решения оптимизационных и обратных задач на графах

Multi-Agent Systems

Мультиагентные системы

1. *В.В. Андреев, С.В. Батищев, К.В. Ивкушкин, Т.В. Искварина, П.О. Скобелев* 233
Инструментальные средства для разработки мультиагентных систем
промышленного масштаба
2. *С.В. Батищев, К.В. Ивкушкин, Т.В. Искварина, В.А. Никифоров, П.О. Скобелев* 241
Анализ возможности применения мультиагентных технологий
в задачах транспортной логистики
3. *Д.В. Абрамов, В.В. Андреев, С.В. Батищев, Т.В. Искварина, Д.В. Карягин,
А.В. Сафронов* 246
Построение онтологий для решения задач транспортной логистики
4. *В.В. Андреев, М.В. Андреев, С.В. Батищев, Т.В. Искварина, П.О. Скобелев* 254
Мультиагентный конструктор и планировщик транспортных сетей
5. *Д.В. Абрамов, К.В. Ивкушкин, Д.В. Карягин, А.В. Олейников, А.В. Сафронов,
П.К. Швейкин* 260
Создание мультиагентной системы поддержки принятия решений
по планированию грузоперевозок в морской логистике
6. *А.В. Алексеев, С.И. Вольман, И.А. Минаков, А.Д. Орлов, М.С. Томин* 270
Создание мультиагентной системы автоматической обработки,
преобразования и коррекции логистических сообщений
стандартных форматов обмена бизнес-данными
7. *В.В. Андреев, О.И. Лахин, И.А. Минаков, А.Н. Сальков, П.О. Скобелев* 277
Развитие элементов самоорганизации и эволюции
в мультиагентном портале социокультурных ресурсов самарской области

- | | | |
|----|---|-----|
| 8. | <i>В.В. Девятков</i>
Мультиагентная архитектура обслуживания запросов к базам данных удаленного восприятия | 282 |
| 9. | <i>С.В. Тиминский</i>
Может ли компьютер быть субъектом права? | 294 |

Regional Management
Региональное управление

- | | | |
|----|--|-----|
| 1. | <i>В.А. Виттих, Г.Д. Светкина, П.О. Скобелев, Е.А. Гриценко, О.Л. Сурнин, Д.В. Волхонцев</i>
Повышение эффективности управления регионом и решение задач ФЦП «Электронная Россия 2002–2010 гг.» с помощью мультиагентных технологий | 300 |
| 2. | <i>Г.И. Гусарова, Г.П. Котельников, В.В. Павлов, В.А. Виттих, А.В. Захарова, Т.В. Тяпухина</i>
Опыт реформирования здравоохранения и пути дальнейшего развития отрасли в условиях административной реформы (на примере Самарской области) | 308 |
| 3. | <i>В.Г. Напреенко, А.С. Нариньяни, Ю.А. Кузин, Н.Е. Яковлев</i>
Оптимизация политики цен и налогов с помощью компьютерной модели региона (на примере Томской области) | 314 |
| 4. | <i>В.А. Виттих, Л.А. Иванова, Е.Н. Королева, Е.Л. Поварова, С.В. Смирнов, Г.Р. Хасаев</i>
Проблемы онтологической спецификации объектов региональной экономики | 322 |
| 5. | <i>Ю.А. Загоруйко, С.В. Булгаков</i>
Использование онтологий для построения инновационных цепочек в системе поддержки инновационной деятельности в регионе | 328 |
| 6. | <i>В.А. Виттих, Г.И. Гусарова, В.В. Павлов, А.В. Захарова, Т.В. Тяпухина, Н.В. Дилигенский, С.В. Кузнецов</i>
Рационализация функций и организационной структуры органов государственной власти в рамках эксперимента в сфере государственного управления в Самарской области (на примере департамента здравоохранения) | 334 |
| 7. | <i>Г.Г. Шалмина, В.В. Тарасевич</i>
Территориальные основы совершенствования системы управления | 340 |
| 8. | <i>В.А. Виттих, Г.Д. Светкина, П.О. Скобелев, Д.В. Волхонцев, Е.А. Гриценко, А.Н. Никитин, О.Л. Сурнин, М.А. Шамашов</i>
Разработка первой очереди системы управления регионом с применением мультиагентных технологий | 346 |
| 9. | <i>С.В. Батищев, В.А. Виттих, Д.В. Волхонцев, Г.И. Гусарова, Г.П. Котельников, С.И. Кузнецов, В.В. Павлов, П.О. Скобелев, О.Л. Сурнин, Л.С. Федосеева, М.А. Шамашов, И.С. Шапиро, Е.В. Чернов</i>
Сетевая мультиагентная модель системы управления здравоохранением региона и система контроля эффективности и качества работы врачей поликлиники | 352 |

- | | | |
|-----|---|-----|
| 10. | <i>В.А. Виттих, П.О. Скобелев, О.И. Лахин, А.Н. Сальков</i>
Мультиагентный интернет-портал для интеграции ресурсов департаментов социального блока самарской области | 362 |
| 11. | <i>А.В. Горохов, В.А. Путилов, А.А. Шестаков</i>
Системная динамика в управлении региональным развитием (на примере Мурманской области) | 370 |
| 12. | <i>Л.С. Федосеева, Н.И. Кантемирова, И.М. Назаркина, Л.А. Щелкунова</i>
Системный подход к управлению качеством оказания медицинской помощи в амбулаторно-поликлинических условиях | 375 |

Control and Measurement in Complex Technical Systems
Управление и измерения в сложных технических системах

- | | | |
|-----|--|-----|
| 1. | <i>Е.Н. Коновалов, В.П. Офицеров, С.В. Смирнов</i>
Повышение эффективности инвестиций в нефтедобыче на основе моделирования | 381 |
| 2. | <i>М.Ю. Лившиц, В.В. Солодьянников</i>
Управление качеством химводоподготовки на ТЭЦ | 386 |
| 3. | <i>В.В. Тарасевич, А.А. Мороз, В.В. Окольников</i>
Математическая модель управления вентиляционным режимом Северо-Муйского тоннеля | 397 |
| 4. | <i>Б.Е. Федун, Т.И. Федунова</i>
Методы оценки эффективности спецификаций алгоритмов бортового интеллекта антропоцентрических объектов | 403 |
| 5. | <i>Н.Ю. Истомина, А.Д. Истомин, А.Н. Жиганов, М.Д. Носков</i>
Информационное обеспечение управления при аварийном выбросе радиоактивных веществ в атмосферу | 410 |
| 6. | <i>А.Г. Кеслер, А.Н. Жиганов, А.Д. Истомин, М.Д. Носков</i>
Программный комплекс для управления разработкой месторождений урана методом подземного выщелачивания | 415 |
| 7. | <i>Ю.Н. Золотухин, А.А. Нестеров</i>
Оптимальный по быстродействию регулятор для произвольного линейного объекта второго порядка | 421 |
| 8. | <i>Г.П. Анишаков, Я.А. Мостовой, А.В. Соллогуб</i>
Устойчивость информационного взаимодействия в сложных технических системах | 425 |
| 9. | <i>В.М. Белкин, В.П. Данильченко, Б.К. Райков, Ю.Н. Секисов, О.П. Скобелев</i>
Экспериментальное подтверждение метода обнаружения срывных и помпажных явлений по колебаниям лопаток | 431 |
| 10. | <i>Л.Б. Беленький, Б.К. Райков, Ю.Н. Секисов, О.П. Скобелев</i>
Одновитковые вихретоковые датчики: от кластерных композиций к кластерным конструкциям | 437 |

11. *С.Ю. Боровик, А.А. Инюцин, О.И. Камаева, В.В. Тулупова* 444
Система измерения многомерных процессов для исследований и диагностики срывных и помпажных явлений в газотурбинном двигателе
12. *С.Ю. Боровик, О.И. Камаева, В.В. Тулупова* 451
Средства моделирования для отладки и проверки работоспособности системы измерения многомерных процессов в газотурбинном двигателе
13. *В.И. Батищев* 457
Априорное моделирование в задачах интерпретации результатов экспериментальных исследований

DESIGNING COMPLEX ENGINEERING SYSTEMS

G. Rzevski

Magenta Corporation Ltd
33 Glasshouse Street, London W1B 5DG
george.rzevski@magenta-technology.com
tel: +44 20 7980 289 626

Keywords: *complex systems, intelligent networks, design of artefacts*

Abstract

Механические изделия типа машин, транспортных средств, космических кораблей и роботов могут разрабатываться не как цельные объекты, а как интеллектуальные сети, что добавит им новое свойство - адаптируемость. Дается два примера применения нового подхода к проектированию технических систем: интеллектуальный компрессор с переменной геометрией и семейство роботов для исследований космоса.

Introduction

Traditionally, we avoid complexity. The aim was always to design well-structured systems, which exhibit predictable behaviour. Artefacts such as large ships, aircraft and process plants have elaborate control systems with explicit feedback loops that ensure that the artefact operates at an equilibrium point. It is considered necessary to ensure that the behaviour is stable, which means that the system, when disturbed, will return to its equilibrium, the sooner the better. The optimal efficiency of the system is tuned to coincide with the equilibrium. Armies, state bureaucracies and large businesses have also been designed as rigid organisations expected to obey commands of the leaders.

There are situations however where stability is undesirable. The classical example is a military aircraft that frequently needs to execute a rapid avoiding manoeuvre. Here the propensity to return to the equilibrium would make any effort to change the operating point more difficult and therefore must be designed out of the system. In fact, whenever the environment is frequently and rapidly changing in an unpredictable manner, the artefact should exhibit Adaptability, ie, the ability to rapidly adjust its behaviour to the dynamics of its environment, rather than stability.

The principle that in a dynamic, unpredictable world *Adaptation* is more important than stability applies equally to physical artefacts and social organisations.

To achieve a really effective adaptation we need to look for help beyond traditional control engineering and management science, which appear to have reached their limits. A good strategy is, of course, to imitate biological systems. To adapt to its environment living systems make use of *Sensory Perception* (detecting and anticipating changes in the environment), *Cognition* (reasoning about perceived changes and deciding on the best action) and *Execution* (controlling the implementation of cognitive decisions) [1]. In situations that require a really fast response, the cognitive stage is short-circuited. Instead of reasoning the system simply selects one of the patterns of behaviour which proved in the past to be effective in similar situations. We tend to call this type of reaction “intuitive”.

Artificial Perception and Cognition systems, particularly advanced ones – like those in humans – are quite difficult to design, the major problems being the collection and formalisation of the essential domain knowledge and mechanism of reasoning under conditions of incomplete, unreliable or frequently changing information. An additional problem is the sheer size of a centralised artificial “brain” and a large number of fast communication channels required for very large artefacts such as jumbo jets or manufacturing plants. A huge intellectual effort has been spent during the last fifty years researching this area, under the name of Artificial Intelligence, albeit with patchy results.

Practical breakthroughs occurred only when the whole problem solving strategy was reviewed and the decision was made to imitate groups rather than individuals – colonies of ants, swarms of bees and, ultimately, human organisations. Note that societies (groups, teams) can be highly adaptive if their organisational structures and cultures are characterised by:

- Significant autonomy of their members, which allows each member to choose whether to co-operate with other members or to compete for limited resources
- Effective communication channels for a rapid exchange of information among all members
- Enforceable rules and/or cultural bonds, which keep the group together in pursuing common goals without unduly limiting autonomy of individuals

For a group to be effective there is no need for individual members to be of exceptional abilities, as long as there are ample opportunities for members to consult each other. For example, groups consisting of members of limited intelligence, like colonies of ants, are capable of quite a remarkable adaptation. This property is given the name Emergence, the creation of intelligent behaviour through the strong interaction among a large number of members rather than through actions of each member. Because of the paramount role of communication links between members of the group, such social structures are known as Networks, and to emphasise the importance of the group intelligence, I refer to them as Intelligent Networks [2].

Examples of social networks include: old-boy networks, drug-dealer networks, terrorist networks and the new breed of virtual enterprises [3].

A social network is a perfect model for the design of the new generation of intelligent software, known as Multi-Agent software, which has a similar configuration (a network with richly interconnected intelligent nodes) and displays adaptability similar to that of flexible societies.

In contrast, social groups with strong hierarchical control, restricted lines of communications between members, and governed by rules and traditions demanding submission to the authority of leaders are known to be rigid and to resist any change even when the change is imposed from the top. Such groups show similarity of structure with conventional software technology, which consists of a hierarchy (rather than a network) of passive software modules.

1 Designing Complexity into Engineering Systems

My contention is that physical artefacts, such as machines, vehicles, spacecraft and plants could be designed as intelligent networks rather than integrated units. By giving them distributed intelligence we make them adaptable to changes in their environments. Let me describe some of the projects in which I was involved that fall into this category.

1.1 Intelligent Compressor

Axial turbo compressors are used in many areas of industry where large quantities of air or gas have to be moved or compressed. Typical examples are in jet engines, large gas turbines, gas-line pumping and in many process plants. All turbo compressors are limited in their performance by the aerodynamic phenomena of stall and surge, where the flow of the gas becomes unstable and can reverse in direction. Stall and surge, if allowed to develop, can cause significant mechanical damage to the compressor.

Present designs of axial compressors use fixed geometry rotor blades and fixed geometry stator vanes, with a limited capability to vary vane angles against pre-set limits, using simple control algorithms. The operating point for the compressor is designed to give an adequate safety margin from the surge line, therefore avoiding the possibility of stall or surge in operation. This surge margin puts a limit on both the work that the compressor can do and its efficiency. For example, large industrial turbo compressors can absorb powers of the order of 50 MW, with annual operating costs of between £5M and £20M, depending on the duty cycle and power charges applied. It is understandable therefore that the costs for even a small drop in efficiency due to a provision for pressure and/or flow with

adequate surge margin, is very significant. The avoidance of surge has important safety implications in aerospace. Surge under extreme manoeuvres has resulted in the loss of aircraft.

Supported by a group of industrial organisations and experienced turbo-machinery designers, the author's team decided to tackle the problem of aerodynamic instability in compressors in a novel way. The decision was made to reconsider the fundamentals of compressor design by removing the usual assumption of fixed or partially variable geometry and to apply concepts from the intelligent network paradigm.

As a result, a design of an axial compressor with variable geometry emerged, where an intelligent agent individually controls each movable element. Agents are then connected into a network and empowered to negotiate among themselves relative positions of movable elements with a view to achieving a performance as close as feasible to the optimum under continuously changing aerodynamic conditions. The overall behaviour of the compressor emerges from the interaction of agents.

The proposed intelligent geometry compressor will operate by using sensors to monitor the aerodynamic conditions at each movable element. Sensor information will be used by local agents, which, through the process of negotiation, will make control decisions and instruct actuators to incrementally adopt the flow path geometry that ensures optimum performance for current aerodynamic conditions.

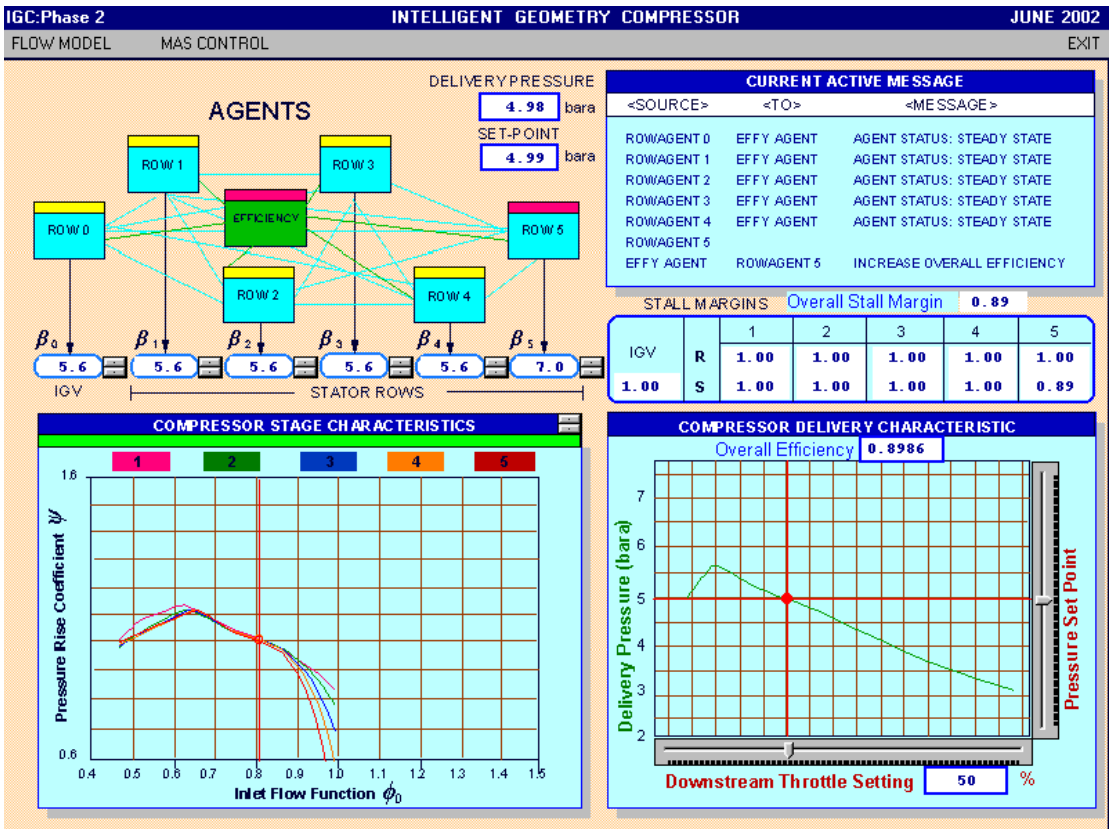


Figure 1 – A multi-agent model of an intelligent geometry compressor

Simulation results indicate that an axial compressor designed as a network of movable elements could be operated over a significantly enlarged envelope without risk of stall or surge [4]. The project has now reached the stage of experimental tests on a variable geometry compressor in Cambridge University Turbomachinery Lab [5].

Utilising embedded processing power it becomes feasible to design into the compressor:

- Self-diagnosing (monitoring compressor conditions and identifying faults when they occur),
- Self-repair by reconfiguration (isolating faulty parts and thus making them harmless)
- Graceful degradation of performance (repositioning remaining healthy parts to achieve a reduced but acceptable level of performance; also, in case of a serious failure of critical elements, such as actuators, agents can revert to the fixed geometry mode of operation)

1.2 A Family of Space Robots

A robot was recently launched by a UK team to explore Mars. Unfortunately, it got lost. It disappeared without a trace among speculations that it landed into a crevice on the Mars surface and lost the communication link with the control station. This disaster could have been avoided by sending to Mars a “network” of smaller robots. In fact, just before the unfortunate Mars robot project started, I joined another team of researchers, involving three universities and two spacecraft manufacturers from the UK, concerned with the development of technologies for the autonomy and robustness in space. Following the design principles described in [2] and briefly outlined in this paper, our team concluded that a family of five much smaller intelligent robots would do a better job under uncertain conditions in space than a single big robot. Each member of the robot family had a limited intelligence and was potentially able to undertake simple tasks such as placing scientific instruments onto a correct location and to provide a variety of services to other members of the family, e.g., cleaning their solar cells if they get covered by the space dust and helping them to get out of cracks in the Mars surface. Multi-agent control system provided intelligence to the robot family without having to rely on “world model” of the Mars environment; just a manageable ontology. The cost and weight per unit performance for the family was estimated to be below the cost and weight of an equivalent single robot. In addition, five smaller robots offered an important advantage in packaging for launch and delivery. The results of this project were ignored with catastrophic consequences.

Conclusion

Only two practical examples out of many available are given here to illustrate the new approach to the design of physical artefacts based on the premise that, for environments characterised by uncertainty, it is advantageous to design complexity into artefacts rather than making them fully predictable.

References

- [1] Rzevski, G (ed), “Mechatronics: Designing Intelligent Machines”, Butterworth Heinemann, 1995.
- [2] Rzevski, G.: “On conceptual Design of Intelligent Mechatronic Systems”. *Mechatronics* 13 (2003) pp. 1029-1044.
- [3] Rzevski, G, Skobelev, P, Batishchev, S, Orlov, A.: “A Framework for Multi-Agent Modelling of Virtual Organisations”. In Camarinha-Matos, L M and Afsarmanesh, H (eds), *Processes and foundations for Virtual Organisations*, Kluwer Academic Publishers, 2003, pp. 253-260.
- [4] Morgan, Gwyn: “Intelligent Geometry Compressor - The Application of Multi-Agent System Technology to the Design of Intelligent Machines”. PhD Thesis, Brunel University, 2002.
- [5] Morgan, G, Rzevski, G., Wiese, P.: “Multi-Agent Control of Variable Geometry Axial Turbo Compressors”. *Journal of Systems and Control Engineering*, issue 13 vol. 218 (2004), pp. 157-171.

БИОИНФОРМАТИКА БАКТЕРИЙ – ОСНОВНЫЕ ЗАДАЧИ И ХАРАКТЕР МОДЕЛЕЙ

Н.А. Кузнецов, М.С. Гельфанд, В.А. Любецкий

Институт проблем передачи информации, Российская академия наук
127994, Москва, Б. Каретный пер, 19
director@iitp.ru, lyubetsk@iitp.ru
тел: +7 (095) 413-46-43, факс +7 (095) 209-05-79

Ключевые слова: биоинформатика, геномика, алгоритм, основные задачи

Abstract

The report includes some essential results of the institute for information transmission problems, Russian Academy of sciences, in the domain of Bioinformatics.

Введение

Развитие методов индустриального секвенирования и связанное с этим быстрое накопление доступных бактериальных геномов делает необходимым развитие методов компьютерной обработки геномных данных. Эта задача поставлена в качестве приоритетной в геномных программах многих стран, в том числе России и США (например, в программе Министерства Энергетики Genomes for Life). В настоящее время доступны методы компьютерной аннотации геномов на основе анализа белковых гомологий и активно развиваются другие методы полногеномного анализа, в частности, выделение позиционных кластеров, филогенетических паттернов и т.п., а также методы метаболической реконструкции. Продолжает развиваться также экспериментальная техника. В частности, уже происходит следующий информационный взрыв в области анализа экспрессии генов на микрочипах.

Одной из весьма существенных областей геномного анализа является анализ регуляции генов. Помимо фундаментального интереса для понимания физиологических процессов, происходящих в бактериальной клетке, такой анализ важен и для смежных областей компьютерной геномики и также имеет большое практическое значение.

С научной точки зрения, интересно не только понимать, *что* делает клетка – ответ на этот вопрос даёт метаболическая реконструкция, но и *когда* и *при каких условиях* она это делает, как включаются и взаимодействуют отдельные метаболические пути и физиологические подсистемы. Выделение регуляторных сигналов – естественное и необходимое продолжение любого эксперимента по массовому анализу экспрессии генов. Более того, без такого анализа результаты подобных экспериментов невозможно интерпретировать, так как отсутствуют другие способы выделить отдельные регуляторные подсистемы и описать регуляторные каскады.

Оказывается, что во многих случаях анализ регуляции даёт возможность уточнить функциональную аннотацию генов, произведенную другими способами. В частности, таким образом часто удаётся уточнить специфичность регуляторных и транспортных белков, а также предсказать новые ферменты, участвующие в рассматриваемом метаболическом пути. Наконец, анализ регуляции совершенно необходим при создании штаммов-гиперпродуцентов или при внедрении в клетку чужеродных генов для создания штаммов с новыми свойствами (например, в целях биологической очистки). Кроме того, многие регуляторные системы являются хорошими потенциальными мишенями для лекарств как широкого, так и узконаправленного действия.

Возникают два типа задач: выделение регуляторного сигнала, то есть общего сигнала, обеспечивающего совместную работу группы генов, иногда в сильно зашумленных данных; и поиск новых сигналов данного вида в геноме, то есть новых генов, регулирующих совместно с заданными. При этом необходимо учитывать, что регуляторные сигналы в нуклеотидных последовательностях бывают двух типов – транскрипционные (то есть отдельные фрагменты ДНК, с которыми связываются белки-регуляторы) и сигналы, связанные с образованием вторичной структуры РНК.

Задача выделения регуляторных сигналов хорошо известна. В Институте проблем передачи информации РАН разработан и запрограммирован вычислительно эффективный алгоритм выделения сигналов. Тестирование и сравнение с доступными программами показало его практическую применимость. Был проведен массовый анализ регуляторных областей в двух важных группах бактерий: гамма-протеобактерии (энтеробактерии-пастереллы-вибрионы) и грамположительные бактерии (бациллы-кlostридии). Для этого были рассмотрены ортологичные ряды генов (отдельно в каждой филогенетической группе), поскольку для таких групп можно предполагать сохранение регуляторного сигнала в родственных геномах.

Проведен сравнительный анализ бактериальных регуляторных систем, важных с практической точки зрения. В ИППИ сформировалась группа ученых, которые в течение нескольких лет ведут работы в указанных направлениях. Многие наши предсказания были подтверждены экспериментально как коллегами, так и независимыми исследователями. Нами исследованы системы утилизации цинка и железа (патогенез), ответа на перенаселенность (стресс и патогенез), ответа на тепловой и гипероксидный шок (стресс), устойчивости к тяжелым металлам (стресс). Работы по двум последним направлениям велись совместно с коллегами из Lawrence Berkeley National Laboratory (USA). При этом рассматривались геномы, секвенированные в Joint Genome Institute DOE USA в рамках проекта Genomes for Life (Министерство Энергетики США). Указанные регуляторные системы особенно важны для анализа бактерий, применяемых в системах очистки.

Методы анализа РНКовых сигналов разработаны существенно слабее. Известно лишь несколько программ поиска известных паттернов, и все они не приспособлены для анализа регуляторных структур. Программы других авторов для выделения консервативных регуляторных структур РНК практически не опубликованы. Нами реализованы два подхода к выделению таких структур. Во-первых, использован тот факт, что регуляторные структуры консервативны в родственных геномах, и поэтому могут быть предсказаны путём выравнивания межгеновых областей ортологичных генов. Версия этого алгоритма реализована и успешно протестирована на модельных ситуациях (например, тРНК). Во-вторых, можно основываться на общих свойствах регуляторных структур – их способности принимать альтернативные конформации, включающие потенциальный терминатор. Этот алгоритм также был реализован и протестирован на ряде известных систем (аттенуаторы аминокислотных оперонов гамма-протеобактерий). При тестировании удалось найти ранее неописанные аттенуаторы в новых геномах. Затем обе эти программы были применены для анализа ортологичных рядов генов из указанных выше и многих других таксономических групп.

Сравнительный подход к РНКовой регуляции был успешно применен нами для анализа нескольких метаболических систем. При этом удалось предсказать либо уточнить функции ряда генов, в основном транспортёров, но также и ферментов. Были предсказаны совершенно новые системы регуляции биосинтеза витаминов (рибофлавина, тиамина и кобаламина), в которых одна и та же альтернативная структура регулирует экспрессию на уровне транскрипции в грамположительных бактериях и трансляции – в грамотрицательных. Впервые была предсказана регуляция за счёт непосредственного связывания конечного продукта метаболического пути со структурным элементом в лидерной области РНК. Недавно эти предсказания были независимо подтверждены в эксперименте. Мы также рассмотрели регуляцию метаболизма ме-

тионина, лизина, ароматических и разветвленных аминокислот у грамположительных и грамотрицательных бактерий.

Помимо анализа регуляции, важную роль в компьютерной геномике прокариот, и в частности, в аннотации генов играет исследование микро- и макроэволюции генов и целых геномов. В первом случае особый интерес представляет анализ гипервариабельных участков генов, поскольку такие участки характерны для поверхностных белков патогенов. Тем самым, учёт гипервариабельных эпитопов важен как для оценки надёжности применяемых диагностических средств и вакцин, так и при создании синтетических вакцин нового поколения. Следует отметить, что такие белки часто оказываются специфичными для узких таксономических групп, и потому слабо поддаются обычному сравнительному анализу.

Нами был значительно усовершенствован известный алгоритм для выделения гипервариабельных участков. Он применен к анализу вирусных геномов.

Наконец, последние две задачи связаны с проблемами макроэволюции. Первой из них является согласование филогенетических деревьев, построенных по отдельным генам, и построение на этой основе общего дерева эволюции видов, учитывающего как неточности в деревьях генов, так и фундаментальные явления дупликации и потери генов. Задача согласования деревьев известна, хотя и далека от удовлетворительного решения. Нами предложен, запрограммирован и протестирован новый алгоритм для решения этой задачи. С его помощью проведен массовый анализ многих деревьев генов, а затем полученные результаты были использованы для решения второй задачи – выделения генов, у которых можно подозревать горизонтальный перенос в ходе их эволюционной истории.

Значимость явления горизонтального переноса в эволюции бактерий стала ясна только в последние два-три года после появления множества полных геномов. Так, массовый горизонтальный перенос, по-видимому, объясняет термофилию ряда бактерий (*Thermotoga*, *Aquifex*, *Thermoanaerobacter*), а также метаболическую пластичность ряда архей (в первую очередь, *Methanosarcina*). Кроме того, следует учесть, что горизонтальный перенос играет важную, если не решающую, роль в таких явлениях, как лекарственная устойчивость и вирулентность патогенов, а также способность бактерий к биодegradации синтетических веществ (фенолы, компоненты нефти) и осаждению тяжелых металлов, что важно для систем биологической очистки.

До сих пор выделение генов, подозрительных на горизонтальный перенос, производилось вручную. Мы создали систему для автоматического выделения таких генов, для чего разработали алгоритм, основанный на следующей идее: горизонтально перенесенные гены – это именно те гены, которые сильнее всего препятствуют вложению дерева генов в дерево видов. Были получены результаты, показавшие применимость и устойчивость этого метода. Осуществлен массовый анализ ортологичных групп генов из базы данных COG, и совместно с коллегами из National Center for Biotechnology Information (National Library of Medicine, NIH, USA) проведен детальный анализ полученных примеров и их систематизация. Помимо вклада в фундаментальную проблему эволюции бактериальных геномов, эти результаты важны для аннотации геномов, описания функций генов. Поэтому мы объединили их с другими результатами геномного анализа, в частности, с результатами выделения регуляторных сигналов. В самом деле, во многих случаях горизонтально перенесенные гены регулируются совместно (например, такая ситуация характерна для систем устойчивости к антибиотикам и тяжелым металлам). Тем самым, такое соединение разных подходов позволяет как уточнить историю соответствующих регуляторных систем, так и быть использованным в практике для борьбы с распространением лекарственной устойчивости у патогенов и при создании новых штаммов для биологической очистки.

В настоящее время стала очевидной необходимость алгоритмов (и компьютерных программ) анализа, не опирающихся на наличие экспериментально охарактеризованной обучающей выборки, т.е. на большое число известных примеров анализируемого явления. До сих пор наличие такой выборки было необходимо для всех доступных алгоритмов анализа бактериаль-

ных геномов; при том, что в большинстве важных случаев адекватная выборка не могла (и не может) быть получена из-за отсутствия достаточного числа надежных примеров, понимания внутренних закономерностей явления и т.д.

В качестве центральных задач геномики прокариот нами рассматриваются и в заметной степени продвинуты решения следующие задачи, направленных на анализ бактериальных геномов.

Основные результаты

Алгоритмы для анализа регуляторных сигналов (1). Разработан сравнительный подход к задаче выявления регуляторных участков в бактериальных геномах. Он с успехом применен при анализе целого ряда регуляторных систем: ответа на стресс (множественной лекарственной устойчивости, теплового шока, SOS-ответа, ответа на перенаселенность), метаболизма (сахаров, аминокислот, витаминов, жирных кислот) и других. В то же время, практика показала недостаточность существующих методов выделения сигнала.

Суть в том, что при анализе геномных данных приходится иметь дело с загрязненными выборками, в которых сигнал присутствует только в части последовательностей из выборки. Существующие алгоритмы, как правило, не в состоянии адекватно анализировать такие данные. Нами предложен и компьютерно реализован существенно новый (очень быстрый по времени работы, примерно квадратичный от размера данных) распараллеливаемый алгоритм, который успешно решает эту задачу для таких выборок (Данилова и др., 2001 – здесь и далее все ссылки относятся к соответствующим годам Activity reports нашего Института; их можно также найти на сайте Института или получить по запросу на адрес director@iitp.ru).

Тестирование этого алгоритма в реальных ситуациях (анализ пуринового и аргининового регулонов, системы катаболитной репрессии) показало его практическую применимость. Проведен массовый анализ регуляторных областей в группе гамма-протеобактерий (энтеробактерии-пастереллы-вибрионы), а также для грамположительных бактерий из группы бациллоклостридий. Для этого рассмотрены выборки ортологичных генов (отдельно из каждой группы), поскольку именно для таких выборок можно предполагать сохранение регуляторного сигнала в родственных геномах.

Отметим, что в обе эти группы бактерий входят как биотехнологически важные организмы (*Escherichia coli*, *Bacillus subtilis*, *Clostridium acetobutylicum*, *Lactococcus lactis*), так и патогены (возбудители чумы, холеры, различных воспалений и кишечных инфекций, сибирской язвы, а также стрептококки, стафилококки, листерии и др.).

Сравнительный анализ регуляторных систем, важных с практической точки зрения (2). Помимо самостоятельной научной ценности, анализ регуляторных систем важен с практической точки зрения, так как он позволяет предсказывать функции ранее неописанных генов, а также прогнозировать ответ клетки на те или иные воздействия. Особый интерес при этом представляют системы патогенности и ответа на стресс. Первый тип систем существенен для медицины, а последняя система – для изучения и улучшения штаммов, используемых при биологической очистки промышленных стоков. Нами исследованы регуляторные системы, среди которых системы утилизации железа и цинка (патогенез), системы ответа на тепловой и гипероксидный шок (стресс), системы ответа на перенаселенность (quorum sensing – стресс и патогенез), системы устойчивости к тяжелым металлам (стресс, крайне важно для систем очистки). Изучена эволюция метаболического пути ответа на гипероксидный шок и распространение систем устойчивости к тяжелым металлам в различных геномах. Также продолжается изучение метаболизма сахаров (в том числе в штаммах, применяемых в пищевой промышленности) и витаминов (для биотехнологических штаммов).

Алгоритмы предсказания регуляторных структур РНК (3). Многие регуляторные системы в бактериальных геномах основаны на образовании вторичных структур РНК. Нами разработаны методы систематического анализа структур такого рода. При этом использованы два типа соображений.

Во-первых, регуляторные структуры консервативны в родственных геномах. Поэтому они могут быть предсказаны путем выравнивания межгенных участков и построения потенциальных вторичных структур РНК для всего выравнивания. Во-вторых, можно использовать общие свойства таких структур – в частности, образование конкурирующих шпилек, из которых одна является регуляторной (терминатором транскрипции или ингибитором инициации трансляции).

В первом случае наши алгоритмы основаны на следующей идее. Сначала в каждом фрагменте выделяются наборы всех (не похожих друг на друга и достаточно мощных) шпилек, затем вычисляется степень похожести любых двух шпилек из двух (любых) разных наборов (в духе алгоритма Смита-Уотермана), к этим шпилькам применяется алгоритм из п. 1 и, наконец, из отобранных таким образом шпилек складываются перекрытия (максимально попарно похожие между собой и не более; вместо перекрытий могут рассматриваться и другие регуляторные вторичные структуры). Этот алгоритм был успешно протестирован на последовательностях тРНК кишечной палочки и некоторых аминокислотных оперонах, а также других вторичных структурах.

Для решения второй задачи разрабатывается алгоритм, основанный на следующей идее. В последовательности РНК ищутся все конфигурации (в частности, пары определенным образом зацепляющихся потенциальных шпилек и т.п.) с качеством, превышающим некоторый порог. Его быстрота обеспечивается новой идеей: сначала находятся определенные численные параметры всех участков последовательности, а по ним строятся сами конфигурации в порядке убывания их качества. Этот алгоритм был успешно протестирован на задаче анализа ряда оперонов аминокислотного синтеза гамма-протеобактерий (гистидина, изолейцина, треонина, триптофана) и затем всех доступных секвенированных бактерий. Соответствующая компьютерная программа была применена, в частности, к ортологичным рядам генов с консервативными терминаторами в лидерных областях.

Сравнительный анализ РНК-овой регуляции (4). Мы применили сравнительный подход для анализа регуляторных систем, основанных на образовании альтернативных структур РНК. Были исследованы аттенуаторы оперонов метаболизма ароматических аминокислот, структуры, регулирующие трансляцию генов рибосомальных белков, RFN, TНI элементы и В12, регулирующие биосинтез витаминов рибофлавина (В2), тиамин (В1) и кобаламина (В12) соответственно и многие другие РНК-овые переключатели. Здесь нами были получены интересные результаты, в частности, был впервые описан механизм регуляции биосинтеза витаминов. В значительной мере мы автоматизировали эту работу, используя программы, созданные в рамках задачи 3.

Анализ гипервариабельных участков генов (5). Приобретение бактериями патогенности в результате горизонтального переноса генов или мутаций, позволяющих занять новую экологическую нишу, часто приводит к появлению мощного селективного давления, стимулирующего отбор новых генетически измененных штаммов. В связи с этим важное практическое значение имеет выявление гипервариабельных участков в генах возбудителей инфекционных болезней, у которых повышенная изменчивость может быть следствием селективного давления иммунной системы организма-хозяина. В данном случае точное знание положения, структуры и степени изменчивости вариабельных эпитопов иммуногенности позволяет, например, судить о надежности применяемых диагностических средств. Подобная информация может оказаться полезной и для разработки синтетических вакцин нового поколения. Наконец, поверхностные белки часто являются специфическими для таксономических групп и потому слабо поддаются сравнительному анализу.

Наглядное представление о распределении консервативных и вариабельных участков гена можно получить при попарном сравнении выравненных нуклеотидных последовательностей и сопоставлении распределения синонимичных и несинонимичных мутаций. Нами значительно усовершенствован алгоритм Nei и Gojobori (USA). Более наглядную картину можно получить при одновременном сравнении нескольких нуклеотидных последовательностей и использовании различных алгоритмов построения графиков генетической вариабельности. Этот подход был апробирован нами на геномах нескольких гепатовирусов (Смирнов Е.В., Зубов В.В., 1997).

В настоящее время доступны геномы многих штаммов целого ряда персистирующих патогенов: стафилококков, стрептококков (гипериммунный ответ на стрептококковую инфекцию приводит к ревматизму), микобактерий (в том числе два штамма возбудителя туберкулеза), хламидий (опубликована гипотеза о том, что хроническая хламидийная инфекция является важным компонентом в патогенезе инфаркта миокарда), малярийных плазмодиев. Мы усовершенствовали разработанную ранее компьютерную программу, предназначенную для оценки распределения мутаций в открытых рамках трансляции.

Согласование филогенетических деревьев (6). Общие алгоритмические приемы, разрабатываемые в рамках предыдущих пунктов 1-4, были применены и для решения более широких эволюционных задач.

Одной из таких задач является реконструкция эволюции видов по молекулярным филогенетическим деревьям отдельных белков.

Деревья белков - очень полезный ресурс, но не вполне адекватный. Во-первых, они не слишком надежны: из-за отсутствия адекватных эволюционных моделей нельзя учесть неравномерности в скоростях эволюции в отдельных таксономических группах, для разных белковых семейств и т.п.

Во-вторых, даже если бы мы получили правильные деревья генов, все равно они по отдельности не дали бы полного представления об эволюции организмов из-за дубликаций генов, их потерь и горизонтальных переносов. Поэтому возникает задача построения деревьев видов на основе согласования деревьев генов и учета некоторых общих представлений об эволюции. Нами построен (Гельфанд и др., 2001; Вьюгин и др., 2001) быстрый алгоритм, который тестировался для групп: позвоночных, основных белков митохондрий эукариот, комплексов ортологических групп из базы данных GenBank (Национального центра биотехнологической информации США). Он основан на следующей идее: сначала выполняются алгебраические вложения каждого дерева генов в некоторое начальное дерево видов, затем оцениваются качества этих вложений и, наконец, происходит последовательная локальная перестройка текущего дерева видов, при которой суммарное качество всех этих вложений убывает вплоть до (локального) минимума. Алгоритм имеет те же очень хорошие вычислительные характеристики, что и в предыдущих пунктах.

Нами построены несколько вариантов филогенетических деревьев для всех основных групп микроорганизмов. На основе оптимального вложения деревьев белков в деревья видов проведена вероятностная реконструкция некоторых событий, происходивших на молекулярном уровне в процессе развития микроорганизмов, в частности оценены возможные дубликации и потери генов (а также их горизонтальный перенос, см. п. 7).

Горизонтальный перенос генов (7). Оказалось, что алгоритм из пункта 6 решает и другую задачу - выделение генов, для которых есть основания подозревать горизонтальный перенос в ходе их эволюционной истории. А именно, это те гены, которые сильнее всего препятствуют вложению данного дерева генов в дерево видов (если разрешенными операциями являются дубликации и потери), т.е. те гены, которые являются, так сказать, артефактами этих вложений. Такой подход основан на том, что ген, перенесенный извне в геном организма на достаточно позднем этапе эволюции, нарушает согласованность между деревом генов, содержащим этот ген, и деревом видов. На основе этого алгоритма получены результаты, которые показали заметную устойчивость выделяемой группы генов при различных вариантах дерева видов. Здесь

кроме совершенствования алгоритма и компьютерной программы была проведена статистическая оценка достоверности выделения такой группы генов нашим алгоритмом (Вьюгин и др., 2001). Разумеется, вопрос о реальном наличии горизонтального переноса может быть решен только биологом-экспертом с учетом всех экспериментальных данных.

Отметим, что недостаточность существующих методов анализа горизонтального переноса хорошо иллюстрируется полемикой, возникшей вокруг компьютерной оценки доли генов человека, которые являются результатом горизонтального переноса из бактериальных геномов (Andersson et al., Science, 2001, 292: 1848; Salzberg et al., Science, 2001, 292: 1903; Ponting, Trends Genet. 2001, 17: 235; International Human Genome Consortium, Nature, 2001; 409: 860). Задача изучения горизонтального переноса генов представляется (кроме её фундаментального значения) весьма важной с практической точки зрения, так как горизонтальный перенос является одним из основных механизмов распространения в бактериальных популяциях лекарственной устойчивости. Кроме того, геномные исследования показывают, что локусы, ответственные за вирулентность (островки патогенности), часто отличаются по своим статистическим характеристикам от других частей генома и также возникают в геномах в результате горизонтального переноса. Наконец, частым объектом горизонтального переноса являются гены устойчивости к тяжелым металлам. В ходе тестирования нашего алгоритма уже был идентифицирован ряд потенциальных горизонтальных переносов. В настоящее время проводится более детальный анализ этих примеров и планируется систематически рассмотреть возможные горизонтальные переносы в семействах генов, представленных в полных геномах.

Массовый анализ регуляторных сигналов по классам организмов (8-9). После решения указанных выше задач 1-7 будет создан необходимый математический и компьютерный аппарат для проведения массового поиска и анализа регуляторных сигналов в классах организмов. Частично он уже создан и применен к классам протеобактерий и грамположительных бактерий, что позволило, с одной стороны, проверить реальные возможности разработанных нами методов и программ, а с другой стороны – послужило делу сравнительного изучения наиболее важных классов бактерий. После завершения этой работы на первый план выходит принципиальной важности задача создания «Рабочего места биоинформатика» - компьютерной среды, в которой будут работать с единообразными типами представления данных программы для всех основных задач геномики прокариот. Предполагается применить наши алгоритмы и программы не только для открытых геномов, но и для ряда геномов, связанных с производством.

Список литературы

- [1] Activity report-2003 of the institute for information transmission problems, Russian Academy of sciences, Moscow, 2004. <http://www.iitp.ru>.

A CONSIDERATION OF PUBLIC CHOICE THEORY FOR THE ANALYSIS OF CHANGE IN PUBLIC POLICY USING MULTI-AGENT SYSTEMS

N.H. Silbert

National IT Research, Inc., Chicago, IL USA
310 Busse Highway, #284 Park Ridge, IL 60068 USA
nsilbert@nationalitresearch.com
tel: +1 (847) 477 7649, fax +1 (847) 655-6120

Keywords: *public choice theory, multi-agent systems, collective action, legislative modeling, James Buchanan*

Abstract

В работе описывается концептуальная модель теории общественного выбора и возможность применения этой модели для анализа изменений в законодательстве и общественной политике. Приводятся характеристики мультиагентной системы, используемой для автоматизации этого анализа.

Introduction

The primary purpose of this article is to explore public choice theory as a conceptual paradigm suitable for automating the analysis of legislative and political change using multi-agent systems. The secondary purpose is to examine the suitability of multi-agent systems to the analysis of political change.

Such a system of analysis, if created, would be very useful in anticipating the legislative, political and economic cost of implementing change in public policy, government and economic regulations. Cost information could drive financial feasibility analysis for change sponsors and provide helpful cost benefit data to compare the results of change with the effort required to effectuate it.

Furthermore, a detailed model of the legislative process provides keen incites into two areas:

Area	Incite
Actors and forces effecting political change	Identification of focus points to support or defend against political change
Process by which legislative change occurs	Identification of potential legislative rules to minimize the unintended consequences vote trading, special interest group capture, etc...

A warning regarding skepticism: there is a strong implication within public choice theory that actors and participants within a democracy seek to further their personal needs and goals and they will use democratic institutions to do so. This theory does not seek to abolish universal values or deny that they exist (such as patriotism, charity, etc...), as it recognizes that many things good for the nation, one's neighbors and society on a whole are also good for the individual; however, it firmly postulates that unless a societal good can also be justified on an individual basis, it may not succeed as a driving political force; e.g. the need for a balanced budget versus the need to increase pension payments for retired individuals – seldom do pensioners vote to reduce their pension payments. Furthermore Public Choice Theory looks at the competing nature of these individual needs in a political marketplace. The consequence is that the theory explores potential failures in the legislative process. Such highlights may be unattractive to those favoring exhaustive government regulation and intervention.

The author hopes that the evaluative model he describes could be considered a modular one and that alternative theories of collective action or public policy analysis could be used to drive an evaluative multi-agent analytic system.

1 James Buchanan and Public Choice Theory

James Buchanan was awarded the Nobel prize in economics in 1986. The theory was formulated in partnership with his intellectual partner, Gordon Tullock. The core elements of Public Choice Theory were jointly envisioned by Buchanan and Tullock between 1958-1959 and authored in 1959 and 1960 in their book: *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Through participation in conferences and establishment of journals, such as *Public Choice* (edited by Tullock), the theory began to build a following through the 1960s. By the 1970s a number of compelling political analysis were implemented and several future government leaders adopted these theories. These theories would then be adapted during the period of great regulatory reform driven by President Reagan in the 1980s, with the famous Gramm-Rudman law its best example. The law limits annual spending by the US Congress and automates budget cuts if the limits are not met.

The key actors in Public Choice Theory are:

- Electorate (citizens)
- Politicians
- Bureaucrats
- Special Interest Groups

In addition, this author would seek to add one other actor or group of actors, traditionally defined as part of special interest groups, these are:

- Large corporations (large institutional, often historic, business enterprises)
- Entrepreneurs (business leaders who rapidly re-assemble and re-create new enterprises, some large, some small)

Whereas there may be risk to the integrity of the theory of public choice by fracturing assumptions of typical actors, in the author's experience, Special Interest Groups often assume a life of their own, independent of the organizations that found them and even continue to support them.

2 Principles of Public Choice Theory

Public Choice theory describes the behaviors of individuals within a political marketplace. It extends the traditional economic model of the marketplace to the public sphere and politics. In a democracy, consensus must be established for legislative change to occur. The establishment of consensus requires exchange of personal, political and economic capital. For example, if a politician wanted the legislature to pass a law, he would make many promises and commitments in order to secure sufficient votes from other politicians for the bill to pass and create the desired law. Fundamentally, these promises and commitments equate to the same economic exchange that occurs in the marketplace.

An important note: to a certain degree Buchanan considers individuals and groups interchangeable; he believes that the same personal interests that drive individual decisions are present in groups within the marketplace and they will act in a manner consistent with their interests. This concept creates the formal connection between the individual in the marketplace and the group in the marketplace. This assumption is also essential to dispel Platonic or Rousseau like notions of a greater public good that transcends individual needs. If individuals cannot avoid pursuing their personal requirements, they will act no differently in political groups.

Within the political marketplace there are many actors. The actors fall into broad or general categories. These categories will often have different if not opposed goals, motivations and strategies. The uniqueness of the different type of strategies helps explain the results of the political process, without which, the results would seem highly irrational. Therefore, it is best to consider the political process to be a competition of different interests each pursuing unique strategies toward unique goals. The ultimate result of the political process is either a blend of different goals built by consensus or the ascendance, at great cost, of one categories goals above all others.

The dynamic of the political marketplace is best understood by examining the various participants (with the author’s extensions to Buchanan’s model):

Actor	Goals
Citizen (electorate)	Government services delivered at little or no cost to the voter.
Politician	Re-election.
Bureaucrat	Increased power through enlarged staff and increased budgets.
Special Interest Groups	Obtain the passage of legislation to favor their particular values or goals.

In addition to classical public choice models, the author believes that other classes of actor are worth considering and lists them here purely for intellectual purposes:

Large Corporations	Seek market expansion and market protection through government involvement
Entrepreneurs	Seek to minimize government regulation to allow rapid and unfettered access to new markets

With such divergent goals it is easy to see how a simple political proposal could find opponents and supporters from the different groups noted above. Furthermore, there are a number of curious interactions and phenomena that occur within and between groups:

- *Voter apathy*

Most citizens believe that their individual vote is one of many and often has little direct impact on the actual process of legislative democracy. Consequently, they decide that there is little to be gained from active review of political activity. Moreover, the act of voting requires significant time spent researching candidates and analyzing their positions, except for the most extreme cases, citizens typically do not feel their vote makes enough of an impact to reward the study required for intelligent voting. Curiously, apathy encourages the birth of special interest groups as the most effective means of political impact by citizens.

- *Bureaucratic Capture*

In her article on Public Choice Theory (The Concise Encyclopedia of Economics), Jane Shaw summarized the economist George Stigler’s observations that Bureaucrats are often compromised by the very parties they are responsible for regulating. Although forbidden, bureaucrats are the indirect beneficiaries of special interest group activity. Interest groups gain political influence through campaign contributions to politicians and can use their influence to ask legislators to increase or decrease the budgets of different regulatory departments. Organizational power becomes very important to Bureaucrats as they do not operate in a profit seeking role. In real world observations, it is not uncommon to see members of industry – typically from the Large Corporations, leave private sector jobs to become Bureaucrats, often regulating the markets of their former employers. Examples can be found in the leadership of the Department of Commerce and the financial markets.

- *Legislative Inefficiency*

Politicians spend the public’s money. This is not the politician’s money, but OPM – *Other People’s Money*. Politicians are not rewarded for the efficiency in which they spend the public’s money, as most voter’s have chosen apathy and do not monitor how the politician administers their trust. Fur-

thermore, pursuing efficiency does not reward politicians. In fact, it is a costly activity which may require they spend their political capital, but for which they would receive no personal reward.

Special interest groups seek to influence politicians' legislative decisions. Their money can be used by the politician in their re-election campaigns. Their role is the farthest thing from voter apathy, for special interests actively monitor and publicize politician's legislative records. Consequently, there is little to protect a legislator's opinion or position from a determined special interest group, except other special interest groups who hold opposite opinions. When special interest groups compete we see as two very determined and focused parties entering a bidding war for legislative favor.

- *Log Rolling / Vote Trading*

Vote trading is a very interesting phenomenon. In a democracy, consensus is required to pass legislation. Vote trading demonstrates the reality that two different voting initiatives do not occur independently in the legislative universe. It further demonstrates that political capital can be accumulated over time into various debts and obligations that can then be exercised at a future date. What is especially amazing, however, is the purely economic nature of political debts. Even the most vociferous of political enemies may collaborate and trade votes on a purely economic basis. By definition, vote trading occurs when two opposing politicians agree to support each other's legislative initiatives due to the critical nature or influence of each other's votes.

Perhaps the fact that politicians spend *Other People's Money* allows them to focus on the political gain of the specific legislation they are seeking to pass, rather than the actual governmental cost of enacting both pieces of legislation, one of which the politician would normally oppose. Vote trading allows each politician to enrich their claims for servicing their constituency when seeking re-election and to effectively ignore the cost of the legislation they passed.

3 A Simulation That Changes Itself

Public Choice theory operates on two different levels. The first, explains how and why political activity follows its own special circuitous path. The second, is a study of self-improvement. Public Choice asks how can a democracy protect itself by adopting rules that limit the negative effects of the democratic process.

Unlike the traditional simulation which models multiple scenarios with different factual data, a public choice simulation will also model changes in the fundamental rules that rule the virtual marketplace within which it exists. Examples might be: legislative spending limits with automatic budget reductions (ala Gramm-Rudman), line item veto granting the President the right to delete specific line items of a bill (many of which are often political favors tacked on to a larger piece of legislation as they would never survive the legislative process on their own merits), term limits for politicians, and decentralization of bureaucratic responsibility (local communities rather than central governments).

To a certain degree a public choice simulation assumes there will be many similar legislative transactions and events occurring over and over again. The point of uniqueness is most likely to be found in the rule changes introduced into the simulation and the effect these rule changes have on overall government spending, growth of government laws and national wealth.

4 The Simulation as a Game

When attempting to describe a simulation environment, one should describe a system where the consequences and activities of one part of the system ultimately affect the rest. Due to the very complex nature of government activity and the need to simplify it for an effective simulation, the author has created a game that will act as a simple model for future multi-agent simulations. This simple model will also help us understand the challenges that a multi-agent simulation will face.

The following are the rules by which a game based on public choice theory may be played:

- 1) Each player chooses one of the standard roles described earlier (more research must be done to clarify and explain the difference between special interest groups, large corporate and entrepreneurial classes). They are now bound by the goals and negotiating strategies of that class.
- 2) Each class has a series of success measurements assigned to it. For example, bureaucrats succeed by increasing the size of the staff or budgets, politicians succeed by buying political popularity, special interests succeed by obtaining the passage of legislation that favors their interest group and the electorate seeks to maximize both societal wealth and free government services.
- 3) Each turn, a card is drawn from a deck labeled legislative opportunities (also called bills). Each opportunity has points (depending on the specific opportunity the points are either positive or negative) assigned to it in several categories:
 - a) Political benefit for re-election (for politicians)
 - b) Changes in regulatory activity (such as more budget for bureaucrats)
 - c) Changes in government services delivered to the electorate (more specifically a targeted group within the electorate such as pensioners)
 - d) Which special interest groups are concerned with this legislation and the economic benefit to their members
 - e) The budgetary cost of the legislation
 - f) The effect of the legislation on the national economy (does it lead to economic growth or contraction for the economy as a whole)
- 4) Each player will either support or attack the legislation. Consensus amongst all players is required to determine the legislation's fate. If the majority of players supports a bill, it passes.
- 5) If a bill passes, then
 - a) Points are awarded to each player according to their class; e.g. politicians receive points for re-election, etc....
 - b) An overall national budget is maintained throughout the game. Each bill that passes will either increase the money in the budget or decrease it.
 - c) Similarly, each piece of legislation effects the national economy, either increasing the gross domestic product (GDP) or decreasing it.
- 6) During the draw of any legislative opportunity, a player may decide if he will cooperate and share the points available to other players of his class (e.g. when a politician vote trades) if the legislation passes, or keep all the points for himself, denying any points to his classmates; e.g. one politician gets all the points, leaving none for his opponents. In this circumstance the consensus process used to determine if a bill passes or not (in step 4) is changed. Each player must vote not only on the bill itself, but for which competitor's version of the bill. Competition can easily occur between different politicians, special interest groups and regulatory agencies (imagine security and foreign trade disagreeing over restrictions on technology exports).
- 7) Every fourth turn a card is also drawn from the structural change card deck. These cards represent significant structural change that dramatically effect game dynamics. For example: some changes may be designed to accelerate the speed at which legislation is passed – for game purposes by instituting a time limit for consensus building and discussion (or perhaps requiring all discussion to be performed in writing using pictures, etc...), while others may have fiscal requirements that either modify bills budgetary cost or cause a potential bill to be rejected and replaced with another. One card example from this deck may be a special interest trump card (for example imagine if green party interests automatically dominated political debate).
- 8) If at any time the GDP falls below a certain level, the economy collapses and all players lose.
- 9) If the various bills deplete the government's budget, then the government is bankrupt and the game ends.

- 10) Otherwise play continues for a proscribed number of turns, perhaps 16. Two bills are considered each year and the game lasts for 8 years.
- 11) Whichever player has the most points at the conclusion of the game, wins.
- 12) Optionally, the players may agree to draw a single card from a third deck called Role Change. This card is drawn at the beginning of the game. Each card in this deck has two pieces of information, the first is the turn at which a role change occurs. The second piece of information is covered and not revealed until the actual turn occurs. It which if not all players changes classes; for example, a politician becomes a bureaucrat in the same way Gov. Ridge of Wisconsin became the head of the department of homeland security or when a bureaucrat leaves government to work for a special interest group. Players keep the points they had accumulated in their previous roles, but they must now score points and follow the strategy of their new class.

Conclusion

The preceding game description can be used as a starting point for the assembly of an ontology of public choice simulation. The game creates a feedback loop through the measurement of government budgets and societal wealth – finite measurements of legislation’s impact on society. Also, the game illustrates that political gain may be achieved by one player by legislation that may in fact hurt the national economy.

The greatest challenge in building a multi-agent simulation will be the evaluation of the impact of legislation on the various classes of participants as well as state budget and national economy. Economic philosophies can far too easily color the cost or benefit of one piece of game legislation over another. Therefore, the arbitrary assessment of the value of legislation is the model’s weakest point.

Consequently, the model may be more effective at analyzing the cost of legislative processes or structural change more so than analyzing the impact of specific pieces of legislation.

КРИТЕРИАЛЬНЫЕ АСПЕКТЫ ПРИНЯТИЯ РЕШЕНИЙ В ЗАДАЧАХ ПРОЕКТИРОВАНИЯ, ОЦЕНКИ ЭФФЕКТИВНОСТИ, ОПТИМИЗАЦИИ СОСТАВА СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ

А.М. Жеребин, В.В. Кропова, В.Г. Баров

ГНЦ РФ ФГУП «Государственный НИИ авиационных систем»

125319, Москва, ул. Викторенко, 7

vvkropova@mail.ru

тел: +7 (095) 157-94-62, факс: +7 (095) 157-53-81

Ключевые слова: функциональная система, внешнее (концептуальное) проектирование, формирование облика, неопределенность действующих факторов, критерии принятия решений, крайняя точка, компромиссное решение, функция индивидуальности

Abstract

Some problems of complex systems design, validation and complement optimization are discussed, such as criterion multiformity, criterion indefiniteness, decisions subjectivity.

1 Критерии принятия решений. Критериальная и другие виды неопределенности

В основе процесса формирования облика и программы создания сложной технической системы (СТС) лежит проведение исследований в таких относительно независимых областях [1, 2] как:

- применение СТС в соответствии с ее целевым назначением в составе функциональной системы (ФС) [3]. Исследуется с использованием комплекса операционных моделей, обеспечивающего оценку эффективности и оптимизацию применения ФС и/или входящих в ее состав технических и организационно-технических систем в операциях различного масштаба и различной интенсивности;
- разработка и производство СТС. Исследуется с использованием комплекса моделей функционирования научно-промышленной базы, позволяющего устанавливать соотношения между характеристиками научно-промышленной базы и объемом и технико-технологическим уровнем выпускаемой продукции;
- поддержание конкурентоспособности СТС рассматриваемого класса на мировом рынке наукоемкой продукции. Исследуется с использованием комплекса операционных моделей, комплекса моделей функционирования научно-промышленной базы, комплекса моделей купли-продажи в условиях кооперации и конкуренции;
- управление развитием функциональной системы. Исследуется с использованием комплекса моделей определения рационального качественного и количественного состава ФС, обеспечивающего оптимизацию этапов жизненных циклов СТС из существующего или исследуемого в качестве потенциально возможного состава ФС. Рациональные планы формируются на основе достижения баланса между требованиями, предъявляемых к СТС по результатам исследований их функционирования в составе ФС, и возможностями выполнения этих требований, устанавливаемых в результате исследований процессов разработки, производства, поддержания функционирования СТС в составе ФС при заданном (прогнозируемом) уровне финансирования.

Принятие наиболее ответственных решений, касающихся облика разрабатываемых СТС, происходит на начальных стадиях проектирования - на этапе так называемого внешнего (концептуального) проектирования – и осуществляется в условиях неопределенности. Можно гово-

речь об эксплуатационной, технологической, финансовой, критериальной и других видах неопределенности.

Под эксплуатационной неопределенностью в данном случае понимается прогнозная неопределенность будущего применения проектируемой СТС в составе ФС, а именно неопределенность характера, объемов, условий выполнения целевых задач.

Технологическая неопределенность обусловлена готовностью научно-промышленной базы к реализации на заданном временном интервале (интервале планирования) совокупности критических технологий, определяющих технико-технологический уровень проекта по созданию перспективной технической системы, и проявляется на этапе внешнего проектирования в отсутствии и/или несвоевременном поступлении достоверной информации о достижимых характеристиках комплектующих различных уровней, непроработанности ключевых технологий, определяющих облик перспективной системы, и др.

Финансовая неопределенность связана с тем, что на этапе внешнего проектирования можно говорить только о прогнозных оценках возможных объема, порядка, источников финансирования проекта по созданию новой техники.

Критериальная неопределенность обусловлена тем, что на этапе внешнего проектирования однозначно не определены те критерии, которыми на различных этапах жизненного цикла проектируемой техники могут руководствоваться лица, принимающие решения (ЛПР), в зависимости от системы предпочтений и складывающейся в каждом конкретном случае ситуации информированности ЛПР.

Можно выделить две группы критериев принятия решений: критерии стратегического управления и критерии рационального выбора тактических решений.

Под стратегическими решениями в данном случае понимаются решения, определяющие динамику качественного и количественного развития ФС. Соответствующие критерии отражают стремление к предельно достижимому

- наращиванию функциональных возможностей ФС (E) в течение периода планирования (T) при прогнозируемом объеме ассигнований ($C_{прогн}$) на содержание и развитие ФС - критерий максимума обобщенной эффективности ($F_1^1 = E \rightarrow \max$ при условии $C \leq C_{прогн}$, E - обобщенная эффективность или производительность ФС, C - ожидаемые расходы на содержание и развитие ФС);
- сокращению дефицита ($E_t_{прогн} - E_t, t = \overline{1, T}$) функциональных возможностей ФС при прогнозируемых по подпериодам планирования ($t = \overline{1, T}$) объемах ассигнований ($C_t_{прогн}, t = \overline{1, T}$) на содержание и развитие ФС - критерий минимума провалов по эффективности ($F_2^1 = F_2^1 \left(\{E_t_{прогн} - E_t\}_1^T \right) \rightarrow \min$ при условии $C_t \leq C_t_{прогн}, t = \overline{1, T}, E_t_{прогн}$ - прогнозируемый уровень требований к функциональным возможностям ФС в t -й подпериод);
- сокращению расходов на содержание и развитие ФС с учетом затрат на программы создания перспективных СТС, предлагаемых для включения в состав ФС, при условии поддержания функциональных возможностей ФС на уровне, соответствующем прогнозируемым потребностям - критерий минимума затрат на содержание и развитие ФС ($F_3^1 = C \rightarrow \min$ при условии либо $E \geq E_{прогн}$, либо $\left(\frac{E_t_{прогн} - E_t}{E_t_{прогн}} \right) \leq \Delta_t, t = \overline{1, T}, \Delta_t$ - допустимый дефицит функциональных возможностей).

К числу наиболее распространенных и часто используемых критериев оптимизации применения СТС в составе ФС (критериев принятия тактических решений) относятся следующие имеющие натуральное или стоимостное выражение критерии:

- критерий минимума суммарной численности (F_1^2) или суммарной стоимости (F_3^2) привлекаемых для участия в планируемой операции ресурсов (сил и средств, в том числе СТС, рассматриваемого класса), при условии выполнения функциональной системой заданных объемов целевых задач ($F_{1,3}^2 \rightarrow \min$ при условии $E \geq E_{\text{прогн}}$);
- критерий минимума суммарной численности (F_2^2) или суммарной стоимости (F_4^2) утрачиваемых в ходе планируемой операции ресурсов при условии выполнения функциональной системой заданных объемов целевых задач ($F_{2,4}^2 \rightarrow \min$ при условии $E \geq E_{\text{прогн}}$);
- критерий минимума суммарного относительного увеличения затрат (F_5^2) на выполнение целевых задач в связи с возможностью реализации нерациональных решений из-за неопределенности действующих факторов при условии выполнения всей (определенной части) номенклатуры целевых задач

$$F_5^2 = \sum_{l=1}^{KT_{\text{ЦЗ}}} \frac{C_{1l} - \min_{x,y(x)} C_{1l}}{\min_{x,y(x)} C_{1l}} \rightarrow \min \text{ при условии } KT_{\text{ЦЗ}} = KT_{\text{прогн}} \left(\frac{KT_{\text{прогн}} - KT_{\text{ЦЗ}}}{KT_{\text{прогн}}} \leq \Delta \right),$$

$KT_{\text{прогн}}, KT_{\text{ЦЗ}}$ - число типов целевых задач, соответственно планируемых для выполнения и выполняемых при располагаемых ресурсах;

C_{1l} - стоимость выполнения одной целевой задачи l -го типа при некоторых фиксированных технических и/или тактических решениях;

$\min_{x,y(x)} C_{1l}$ - минимальная стоимость выполнения одной целевой задачи l -го типа, т.е. стоимость выполнения одной целевой задачи для оптимального вектора технических и/или тактических решений;

Δ - допустимый процент невыполнения целевых задач из заданной номенклатуры.

- критерий минимума наибольшего относительного увеличения затрат (F_6^2) на выполнение целевых задач в связи с возможностью реализации нерациональных решений из-за неопределенности действующих факторов при условии выполнения всей (определенной части) номенклатуры целевых задач

$$F_6^2 = \max_{l=1, KT_{\text{ЦЗ}}} \frac{C_{1l} - \min_{x,y(x)} C_{1l}}{\min_{x,y(x)} C_{1l}} \rightarrow \min \text{ при условии } KT_{\text{ЦЗ}} = KT_{\text{прогн}} \left(\frac{KT_{\text{прогн}} - KT_{\text{ЦЗ}}}{KT_{\text{прогн}}} \leq \Delta \right).$$

2 Преодоление неопределенности действующих факторов

Учитывая неопределенность на этапе внешнего (концептуального) проектирования основных действующих факторов, целесообразным представляется подход, при котором задача выбора рационального облика СТС решается по схеме: сначала формируется множество так называемых крайних точек, а затем с использованием формальных и/или неформальных процедур достижения компромисса на множестве крайних точек определяется облик проектируемой системы.

2.1 Формирование множества крайних точек

Крайняя точка – это вариант облика проектируемой системы, полученный в результате решения оптимизационной задачи, формальная постановка которой имеет ограниченный (локальный) характер в силу следующих обстоятельств. Все действующие факторы разбиваются на две группы, причем в первую группу могут быть включены только те факторы, которые характеризуются прогнозной неопределенностью. Значения факторов из первой группы фиксируются, тем самым снимается соответствующая неопределенность. Для факторов второй группы допускается варьирование их значений в пределах, определяемых зафиксированными значениями факторов первой группы.

Для определенности процедуру формирования крайних точек рассмотрим применительно к ситуации, когда первую группу факторов составляют критерии рационального выбора, которыми в принципе могут пользоваться разработчики и эксплуатанты проектируемой системы.

Крайняя точка (вектор обликовых параметров проектируемой системы), соответствующая сложившейся в силу принятых допущений ситуации информированности, отыскивается в соответствии с правилами.

$$(1) \quad x^* = \arg \min_{x \in X_{\text{дон}}} F_{\text{разр}}(x, y^*(x)),$$

$$(2) \quad y^*(x) = \arg \min_{y \in Y_{\text{дон}}(x)} F_{\text{экспл}}(x, y),$$

$$(3) \quad X_{\text{дон}} = \{x : f(x) \leq 0\},$$

$$(4) \quad Y_{\text{дон}}(x) = \{y(x) : g(x, y) \leq 0\},$$

$x, y(x)$ - вектора технических и тактических решений, принимаемых на этапах соответственно разработки и применения проектируемой системы;

$x^*, y^*(x)$ - рациональные при данной постановке задачи значения векторов технических и тактических решений;

$X_{\text{дон}}, Y_{\text{дон}}(x)$ - области допустимых технических и тактических решений на этапах соответственно разработки и применения проектируемой системы;

$f(x), g(x, y)$ - ограничения (вектора ограничений) на технические и тактические решения на этапах соответственно разработки и применения проектируемой системы;

$F_{\text{разр}}(x, y), F_{\text{экспл}}(x, y)$ - критерии фиксированного для данной постановки задачи вида из числа критериев рационального выбора, которыми потенциально могут воспользоваться разработчик ($F_{\text{разр}}$) и эксплуатант ($F_{\text{экспл}}$) проектируемой системы.

В результате решения серии оптимизационных задач вида (1)–(4) формируются множества крайних точек, на котором отыскивается компромиссное решение по выбору облика проектируемой системы.

2.2 Формирование компромиссного решения на множестве крайних точек

Сведем совокупность факторов, связанных со всеми видами неопределенности кроме критериальной, к множеству расчетных вариантов. Другими словами, при формировании облика СТС будем рассматривать множество расчетных вариантов и множество критериев рационального выбора.

Итак, рассматриваются

- 1) Пространство принимаемых решений (пространство выбора):

$$R = \{r = (x, y), x \in X, y \in Y\},$$

$$X = \{x = (x_1, x_2, \dots, x_m)\} \left\{ \begin{array}{l} \text{- множество предлагаемых разработчиками альтернативных вариантов облика СТС (задаваемое множество), если реализуется подход на основе сравнения альтернативных вариантов СТС;} \\ \text{- множество синтезируемых в процессе моделирования вариантов облика СТС (формируемое множество), если реализуется подход на основе непосредственного синтеза обликовых параметров СТС;} \end{array} \right.$$

$$Y = \{y = (y_1, y_2, \dots, y_n)\} \quad \text{- множество тактических решений СТС.}$$

2) Пространство условий принятия решений (пространство условий выбора)

$$\Omega = \{(z, C), z \in Z, C \in W\},$$

$Z = \{z\}$ - множество расчетных вариантов при формировании облика СТС;

$W = \{C\}$ - множество критериев рационального выбора.

Для определенности будем считать, что в качестве определяющих рассматриваются два фактора: условия применения СТС и условия реализации СТС, т.е.

$$z = (u, v), u \in U, v \in V$$

$U = \{u = (u_1, u_2, \dots, u_l)\}$ - множество условий применения СТС (тип операций, характер и объем целевых задач и др.);

$V = \{v = (v_1, v_2, \dots, v_k)\}$ - множество условий реализации СТС или множество вариантов реализации критических технологий.

Процедура принятия компромиссного решения состоит в следующем.

3) Снятие неопределенности по условиям применения и реализации (если задан порядок расчетные варианты - критерии).

$$x_{\text{рац}}^* \in \left\{ \begin{array}{l} \text{либо } \left\{ x \mid (\forall z, z \in Z), \left(\frac{\Delta C(x, y, z)}{C(x_{\text{рац}}, y_{\text{рац}}, z)} \leq \delta \right) \right\}, C \in W, \text{ если имеет место} \\ \text{ситуация, представленная на рис. 1.а} \\ \text{либо } \left\{ x \mid x = \arg \min_{x, y} \left(\alpha_1 \sum_z \frac{\Delta C(x, y, z)}{C(x_{\text{рац}}, y_{\text{рац}}, z)} + \alpha_2 \max_z \frac{\Delta C(x, y, z)}{C(x_{\text{рац}}, y_{\text{рац}}, z)} \right) \right\}, C \in W, \\ \text{если имеет место ситуация, представленная на рис. 1.б} \end{array} \right.$$

$$\Delta C = C(x, y, z) - C(x_{\text{рац}}, y_{\text{рац}}, z)$$

$$\langle x_{\text{рац}}, y_{\text{рац}} \rangle = \arg \min_{x, y} C(x, y, z)$$

$C(x, y, z)$ - используемый в данном конкретном случае критерий рационального выбора из заданной номенклатуры критериев;

$x_{\text{рац}}, y_{\text{рац}}$ - рациональные значения векторов технических и тактических решений СТС.

4) Снятие критериальной неопределенности (если задан порядок расчетные варианты - критерии).

$$x_{\text{рац}}^{**} \in \left\{ \begin{array}{l} \text{либо } \left\{ x \mid x \in \bigcap_{C \in W} \{x_{\text{рац}}^*(C)\} \right\}, \text{ если } \bigcap_{C \in W} \{x_{\text{рац}}^*(C)\} \neq \emptyset, \text{ т.е. имеет место} \\ \text{ситуация, аналогичная представленной на рис. 1.а} \\ \text{либо } \left\{ x \mid x = \arg \min_{x, y} \left(\alpha_1 \sum_C \frac{\Delta C(x, y, z)}{C(x_{\text{рац}}^*, y_{\text{рац}}^*, z)} + \alpha_2 \max_C \frac{\Delta C(x, y, z)}{C(x_{\text{рац}}^*, y_{\text{рац}}^*, z)} \right) \right\}, \\ \text{если имеет место ситуация, аналогичная представленной на рис. 1.б} \end{array} \right.$$

Выражение $\alpha_1 \sum_C \frac{\Delta C(x, y, z)}{C(x_{рац}, y_{рац}, z)} + \alpha_2 \max_C \frac{\Delta C(x, y, z)}{C(x_{рац}, y_{рац}, z)}$ является линейной сверткой двух видов критериев, которые отражают две крайние позиции в принятии решений: результат в среднем $\sum_C \frac{\Delta C(x, y, z)}{C(x_{рац}, y_{рац}, z)}$ и гарантированный результат $\max_C \frac{\Delta C(x, y, z)}{C(x_{рац}, y_{рац}, z)}$. Значения коэффициентов α_1 и α_2 либо принимаются по умолчанию в соответствии с ранее принятыми правилами, либо устанавливаются непосредственно исследователем, экспертом, лицом, формирующим или принимающим решения, и фактически определяют структуру критерия, по которому принимается компромиссное решение.

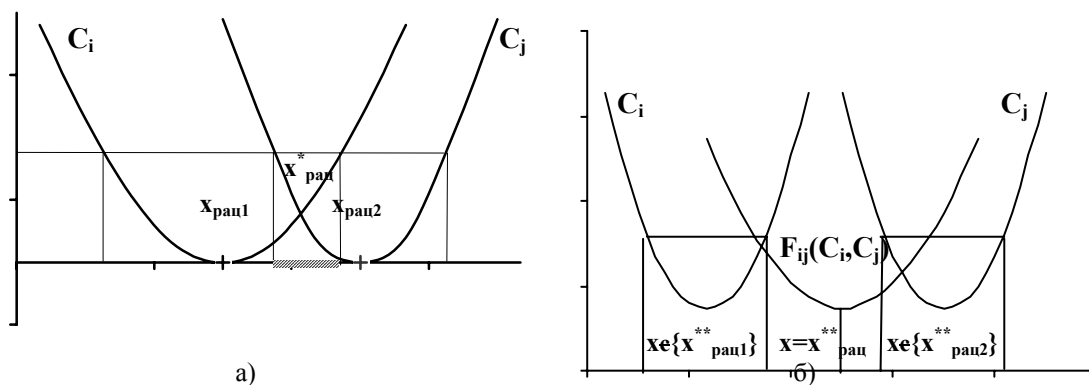


Рисунок 1 – Ситуации при принятии компромиссного решения.

3 Учет фактора субъективности при принятии решений

Состояние выбора (принятия решения) складывается из четырех компонент:

- 1) субъект, производящий выбор A : целеустремленный индивид или система;
- 2) окружение выбора $S \in \{S_k, k = \overline{1, K}\}$: функциональное окружение индивида или системы, проявляющих выбор;
- 3) доступные способы действий $C_i, i = \overline{1, I}$: структурно различные типы поведения субъекта в структурно определенном окружении;
- 4) возможные при таком окружении результаты $Q_j, j = \overline{1, J}$.

Существенные связи между этими компонентами полностью отражаются в трех типах мер, являющихся параметрами целеустремленного состояния:

- вероятность выбора (P_i) - вероятность того, что субъект A будет продуцировать способ действия (C_i) в окружении выбора S

$$P_i = P\{C_i | A \in S\}, i = \overline{1, I};$$

- эффективность способа действий (E_{ij}) - вероятность того, что некоторый способ действия C_i будет продуцировать определенный результат Q_j в определенном окружении S , если данный субъект A выберет именно его

$$E_{ij} = P\{Q_j | \text{выбирает } C \in S\}, i = \overline{1, I}, j = \overline{1, J};$$

- удельная ценность (V_j) результата (Q_j) или степень стремления (DI_j) субъекта A к результату (Q_j) относительно полного множества взаимоисключающих результатов (Q_j) в ситуации стремления в окружении выбора S - вероятность того, что A выберет способ действий, имеющий максимальную эффективность по Q_j .

Поскольку способы действий и результаты зависят как от окружения, так и от субъекта, индивидуальность субъекта, его личность определяется тем, как его вероятности выбора (P_i), эффективности (E_{ij}) и удельные ценности (V_j) зависят от свойств ситуации.

$$\begin{aligned} P_i &= f[\{C_i\}, \{Q_j\}, S_k], \\ E_{ij} &= g[\{C_i\}, \{Q_j\}, S_k], \\ V_j &= h[\{C_i\}, \{Q_j\}, S_k]. \end{aligned}$$

Другими словами, если два индивида находятся в одной и той же ситуации выбора, разница между ними должна проявляться в значениях их вероятностей выбора, эффективностей и удельных ценностей.

Сумма произведений этих трех параметров равна для данного субъекта ожидаемой удельной ценности (EV)

$$EV = \sum_i \sum_j P_i \cdot E_{ij} \cdot V_j.$$

Поскольку величины P_i, E_{ij}, V_j являются функциями ситуации выбора, то и ожидаемая удельная ценность будет функцией ситуации выбора

$$EV = \pi[\{C_i\}, \{Q_j\}, S_k]$$

В таком случае личность (индивидуальность) субъекта есть функция π , связывающая ожидаемую удельную ценность в любой ситуации выбора со свойствами возможных способов действий, их возможных результатов и существенными переменными окружения. Т.е. личность (индивидуальность) есть наблюдаемая функция, показывающая, как индивид или система преобразует параметры выбранной ситуации в ожидаемую удельную ценность (см. рисунок 2).

Задачу построения функции индивидуальности можно разбить на три относительно независимых типа исследований в зависимости от того, какая берется зависимая переменная:

- вероятность выбора - изучение привычности (предпочтения средств);
- эффективность выбора - изучение знания (чувствительности к эффективности);
- удельная ценность - изучение стремления (предпочтения целей).

Итак, функцию личности можно получить, если известны

- функция привычности (f_F) для субъекта способа действия (C_i) относительно результата (Q_j) и доступного множества способов действий $\{C_i\}$ - это функция, связывающие вероятность выбора с другими параметрами ситуации выбора, т.е. удовлетворяющая соотношению

$$P_i = f_F[\{C_i\}, V_i, L(E_i)],$$

где $L(E_i)$ - уровень эффективности способов действий по результату (Q_j), причем

$$0 \leq L(E_i) \leq 1, \quad \sum_i L(E_i) = 1;$$

- функция знания (f_k) субъекта способа действия (C_i) относительно результата (Q_j) в окружении S - это функция, связывающая эффективность выбора с другими параметрами ситуации выбора, т.е. удовлетворяющая соотношению

$$DK_{ij} = f_k(V_j|S)$$

где DK_{ij} j степень знания, которым субъект располагает о способе действия (C_i) относительно предпочтительного результата с удельной ценностью $V_j = 1$ в окружении выбора S .

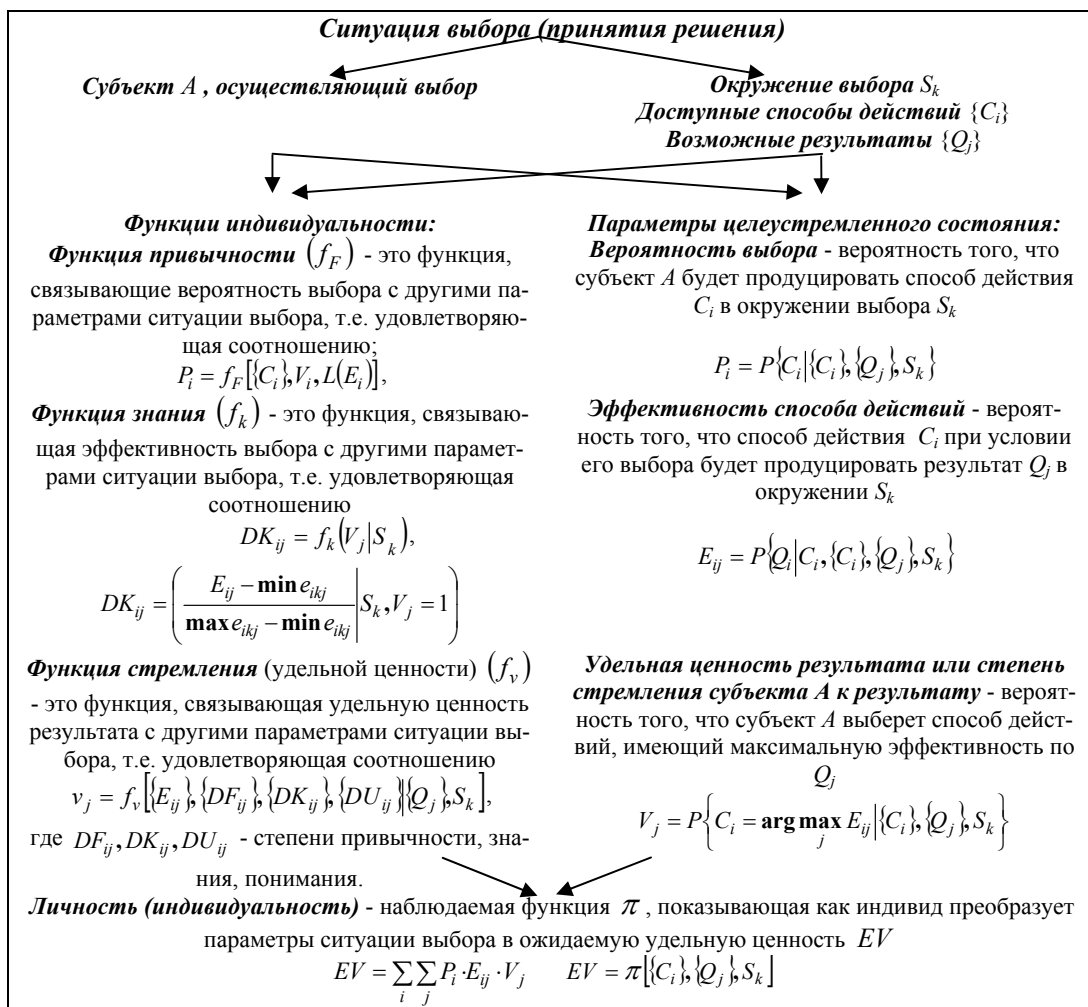


Рисунок 2 – Характеристики личности (индивидуальности) субъекта.

Это степень знания задается формулой

$$DK_{ij} = \left(\frac{E_{ij} - \min e_{ikj}}{\max e_{ikj} - \min e_{ikj}} \right) \Big|_{S, V_j = 1},$$

где $\min e_{ikj}, \max e_{ikj}$ - наименьшее и наибольшее возможные эффективности C_i по результату Q_j в S .

Степень знания способа действий относительно какой-либо цели в фиксированном окружении является мерой управляющих возможностей субъекта для получения данного результата относительно максимально возможных управляющих воздействий;

- функция стремления (удельной ценности) (f_v) субъекта к результату (Q_j) , относительно полного множества взаимоисключающих результатов $\{Q_j\}$ в ситуации стремления в окружении выбора S - это функция, связывающая удельную ценность результата с другими параметрами ситуации выбора, т.е. удовлетворяющая соотношению

$$v_j = f_v[\{E_{ij}\}, \{DF_{ij}\}, \{DK_{ij}\}, \{DU_{ij}\}, \{Q_j\}, S],$$

где $DF_{ij}, DK_{ij}, DU_{ij}$ j -степени привычности, знания, понимания.

В качестве меры степени переменчивости индивида по отношению к некоторому результату может рассматриваться производная от функции стремления индивида к этому результату по эффективности данного способа действий в его продуцировании.

Знание о функциях индивидуальности ЛПР может использоваться при формировании интегрального критерия принятия компромиссных решений относительно облика проектируемой СТС, а именно, при назначении коэффициентов важности, с которыми критерии из числа рассматриваемых критериев учитываются в интегральном критерии.

Список литературы

- [1] Жеребин А.М., Кропова В.В. Методологические основы моделирования и управления развитием функциональных (целенаправленных) систем // Проблемы управления и моделирования в сложных системах: Труды II Международной конференции. – Самара: СНЦ РАН, 2000. С. 80–85.
- [2] Жеребин А.М., Кропова В.В., Баров В.Г. Интегрированная среда моделирования для прогнозирования направлений возможного использования научно-технологических разработок при создании перспективных технических систем // Проблемы управления и моделирования в сложных системах: Труды V Международной конференции. – Самара: СНЦ РАН, 2003. С. 15–21.
- [3] Жеребин А.М., Кропова В.В., Мартиросов С.Р. Подход к моделированию в интересах стратегического управления и исследования устойчивости функциональных систем // Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции. – Самара: СНЦ РАН, 2002. С. 31–39.

CONTEXT-BASED KNOWLEDGE LOGISTICS FOR DISASTER MANAGEMENT

A.V. Smirnov, M.P. Pashkin, N.V. Chilov, T.V. Levashova, A.A. Krizhanovsky

St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences

39, 14th Line, St.Petersburg, 199178, Russia

smir@mail.iias.spb.su

тел: +7 (812) 328-8071, факс: +7(812) 328-0685

Keywords: *knowledge management, logistics, disaster aid, web-service, distributed knowledge sources, knowledge integration*

Abstract

Системы транспортировки играют существенную роль в решении проблем, возникающих в случае природных катаклизмов, стихийных бедствий и катастроф. Такие современные технологии как, например, Интернет делают возможным создание интеллектуальных систем транспортировки для расширения их возможностей для своевременного реагирования на эти катастрофы. Эта статья представляет подход к реализации Интернет сервисов для прокладки маршрутов в больших системах транспортировки, используя преимущества идеи логистики знаний. Логистика знаний основана синергетической интеграции знаний, приобретенных из распределенных источников необходимых для получения нового или дополнения существующего знания. Данная статья показывает масштабируемость и расширяемость представленного подхода при добавлении новых источников знаний.

Introduction

Nowadays organizations have to restructure in order to respond to their rapidly changing environment. This causes a growing recognition of the need for new kinds of organizational structures that could be distributed, mobile and flexible and therefore exhibit characteristics of innovation, resilience, and self-management.

One of the most widespread forms of organizational structures the form of networked organization can be mentioned. Among the main advantages of the networked organization the following ones are selected:

- scalability: new members can join a networked organization by establishing relationships with some of its members,
- robustness: a networked organization will continue to operate even if some of the network nodes stop,
- sensitivity and adaptability: when flexible enough the relationships between the organization's units (the network's arcs) may be easily and quickly readjusted in accordance with changes in the environment,
- intensive knowledge/skills/experience exchange between the organization's units resulting in maximization of the knowledge potential of the organization.

As it can be seen new forms of organizational structures including networked and virtual organization rely on intensive usage of knowledge and therefore on modern information technologies dealing with data and knowledge. Currently, knowledge often referred to as corporate knowledge in the contexts of the business and manufacturing is the key resource in the modern era of information.

We will proceed with the following definition, taken from [1]:

Context: any information that can be used to characterize the situation of entities (i.e. whether a person, place or object) that are considered relevant to the interaction between a user and an application,

including the user and the application themselves. Context is typically the location, identity and state of people, groups and computational and physical objects.

Growing importance of intelligent support of the knowledge customers causes a need for acquisition, integration, and transfer of the right knowledge from distributed sources located in an information environment. This knowledge has to be delivered in the right context to the right person, in the right time for the right purpose. These activities called *knowledge logistics* are required for global awareness, dynamic planning and global information exchange in the information environment.

Transportation systems play an important role in coalition-based application implemented as networked organizations. An intelligent transportation system may significantly enhance the transportation system abilities. Therefore the paper proposes an application of knowledge logistics based on the concept of open services in a distributed environment as an intelligent service for creation efficient routing plans (as one of the major tasks of transportation system management for responding disasters and organizing evacuations) under given constraints and preferences.

1 KSNet-Approach for Transportation System Management

As it is mentioned above, the being described approach is devoted to the knowledge logistics problem. It considers this problem as a problem of a Knowledge Source Network (KSNet) configuration, in this connection it has been referred to as KSNet-approach.

1.1 KSNet-Approach to Knowledge Logistics

Knowledge logistics assumes dealing with knowledge contained in distributed and heterogeneous KSs. As a result of this, the approach is oriented to ontological model providing a common way of knowledge representation. KSNet-approach proposes ontology-driven methodology to knowledge source network configuration. This approach is described in detail in [2].

Knowledge sources (KSs) to be used comprise end-users / customers, loosely coupled knowledge sources / resources, and a set of tools and methods for information / knowledge processing. In the context of the paper the configured KSNet is thought of as a networked organization, where the listed above constituents of the KSNet correspond to the organization nodes.

Because of distributed structure of networked organization, its behavior of open system, and orientation on the Internet as the e-business environment the technology of Web-services for the approach implementation has been applied.

As a formal model for knowledge integration the ontology model with the knowledge representation formalism of object-oriented constraint networks was chosen. This allows simplifying the formulation and interpretation of real-world problems which in the areas of engineering, manufacturing, management, etc. are usually presented as constraint satisfaction problems (e.g., [2, 3]). As a constraint-based tool ILOG tool has been chosen [5]. The system implementing the approach inherited its name; it is referred to as the system "KSNet". The detailed description of the multiagent system architecture and its functionalities can be found in [4].

1.2 Health Service Logistics in Coalition Operations Other than War

Coalition operations other than war include a wide range of activities involving different people and organizations. In classical war operations the technology of control is strictly hierarchical, unlike operations other than war are very likely to be based on cooperation of a number of different, quasi-volunteered, vaguely organized groups of people, non-government organizations, institutions providing humanitarian aid but also army troops and official governmental initiatives. Here many participants will be ready to share information with some well specified community [6].

All joint doctrine and tactics, techniques, and procedures are organized into a comprehensive hierarchy as shown in (Figure 1). As a case study for the project the area of logistics was chosen because logistics problems can be widely used in a number of areas (e.g., other than war operations, supply chain management, transportation systems, etc.) and it is of a high importance for completion

of joint missions [7]. Focused logistics operations and/or Web-enhanced logistics operations address sustainment, transportation and end-to-end rapid supply to the final destination. Here the distributed information management and real-time information fusion to support continuous information integration of all participants of the operations are needed [8].

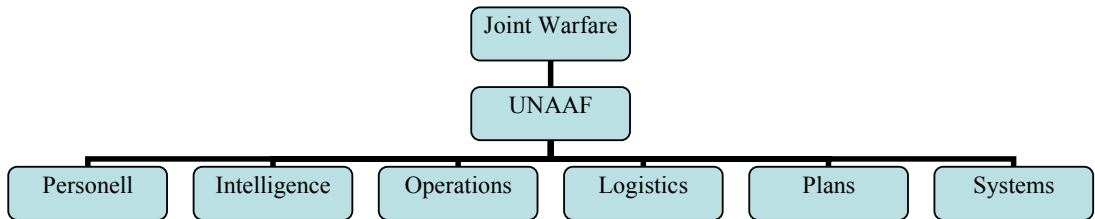


Figure 1 – Joint doctrine hierarchy (adapted from [7]).

Logistics support in coalition operations presents numerous challenges due to a variety of different policies, procedures and practices of the members of the operations, e.g., difference in doctrine, logistics mobility, resource limitations, interoperability concerns, and competition between participants for common support.

In [7], six major principles of joint activities logistics applying to operations other than war are selected. They include:

- *Objective.* There must be a clearly defined, decisive and attainable objective, and all the efforts of each operation member have to be integrated into the total effort of achieving strategic aims and cumulating in the desired end state.
- *Unity of effort.* There must be a close coordination of all the operation members provided leading toward the main goal and every subgoal.
- *Legitimacy.* Legitimacy involves sustaining the people's willingness to accept the right of the operation' leader to make and carry out decisions so that their activities would complement, not detract from, the legitimate authority of the leaders.
- *Perseverance.* In coalition operations strategic goals may be accomplished by long-term involvement, plans, and programs. Short duration operations may occur, but these operations have to be viewed as to their impact on the long-term strategic goals.
- *Restrain.* Coalition other than war operations put constraints on potential actions that can be undertaken by the operation's members to achieve their goals.
- *Security.* Security is a very important issue in coalition operations, especially in those related to healthcare and involving military forces. The operation's leaders and members have to ensure that they include security measures.

Coalition other than war operations may have different missions. E.g., they can be related to disaster relief, noncombatant evacuation, humanitarian assistance, peace operations, and other. For the project a task of disaster relief operation form the area of health logistics has been chosen; particularly, it is devoted to mobile hospital configuration.

1.3 Knowledge Sources for Creation Routing Plans

For the task of the routing plan creation the following knowledge sources can be considered: available suppliers (constraints on suppliers' capabilities, capacities, locations); available providers of transportation services (constraints on available types, routes, and time of delivery); factors influencing on route availabilities as the geography and weather of the considered region (constraints on types, routes, and time of delivery, e.g. by air, by trucks, by off-road vehicles).

The problem of an automatic knowledge seeking is a future research. For the case study a list of KSs containing information for the user request processing was prepared by an expert team.

In the presented case study a fictitious Binni region [9] is considered. The aim of the used Binni scenario is to provide a rich environment, focusing on new aspects of coalition problems and new technologies demonstrating the ability of distributed systems for intelligent support to supply services in an increasingly dynamic environment. The transportation system is presented by a weighted graph of routes assigned with transportation time, costs and route type (highway, ground road, transportation by a plane, etc.). This information is stored in an external database accessed remotely.

Each supplier is characterized by its location, capabilities (which products it can produce), capacities (how many products it can produce by a certain moment), schedule (when it products can be picked up).

In the presented example the one-customer model is considered. The reason for this is that the considered transportation task is a part of a bigger task of a portable hospital configuration described in detail in [10].

In order to provide up-to-date transportation plans the system monitors the current situation in the region. For this purpose an emulated news Web site has been implemented (Figure 2). A specially designed wrapper reads news and finds which cities/areas are not currently available for transportation. Besides, it reads weather conditions and accordingly corrects transportation time and costs for appropriate routes.

This list of sources is not fixed. The scalable architecture of the system KSNNet allows seamless attaching of new sources in order to get new features and to take into account more factors for tasks solved.



Figure 2 – Emulated news Web-site.

2 Knowledge Logistics as an Intelligent Web-Service

Open services is an emerging technology driven by the will to securely expose business logic beyond the firewall. Through open services organizations can encapsulate existing processes, publish them as services, search for and subscribe to other services, and exchange information throughout and beyond the enterprise. Open services will enable application-to-application e-marketplace interaction, removing the inefficiencies of human intervention.

As it was mentioned a networked organization environment is characterized as an open service system (i) in which users, units and other resources come and go on a continual basis, and (ii) in which entities provide services to each other under various forms of a contract (or agreement). The key components of service oriented architecture are as follows: *service owners* that offer *services* to *service consumers* under particular *contracts*. Each owner-consumer interaction takes place in a given *marketplace* whose *rules* are set by *market owner*. The service owners and service providers interact with each other in a particular *environment context*. The open service oriented model applied to the system "KSNet" looks as shown in (Figure 3). The system acts as a service provider for knowledge customer services and at the same time as a service requestor for knowledge suppliers.

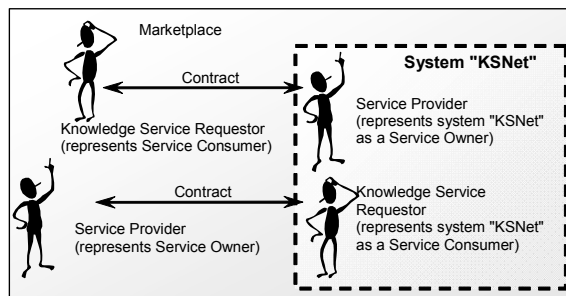


Figure 3 – Open service oriented model of networked organization.

The conceptual scenario of Web services is shown in (Figure 4). A service requester sends a request for an appropriate service ("Find Service") to the UDDI Registry. The Registry returns a reference to an appropriate service provider to the requestor. Then, the requestor sends a request to the service and receives the result from it.

In the system "KSNet" the Web-based open service model is used. The proposed scenario is presented in (Figure 5). The main specific is that the service passes the request into the system where it goes through all the stages of the request processing scenario. When an answer for the request is found it gets to the service and then it is passed to the requestor.

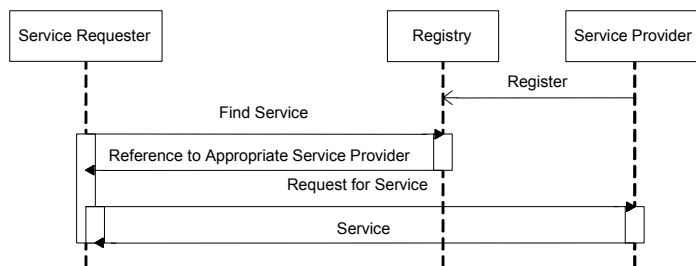


Figure 4 – Conceptual Web service scenario.

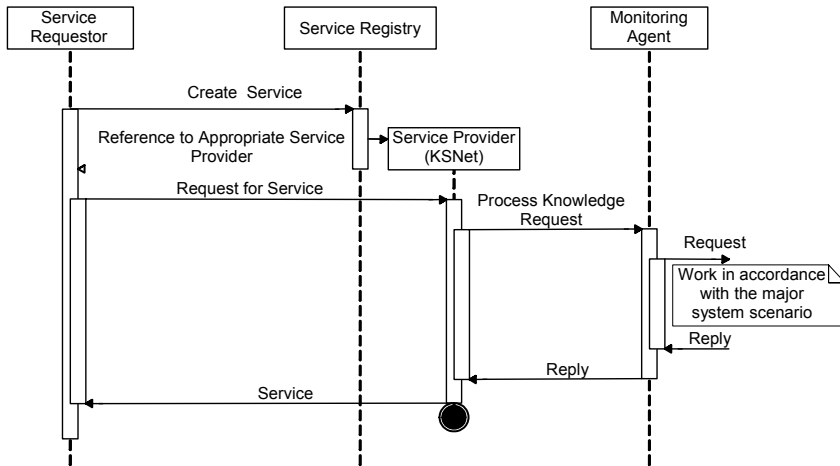


Figure 5 – Web service scenario for the system "KSNet".

3 Implementation and a Case Study

The user represented by a Web-service client enters a request into the system via the developed SOAP- based interface. It has been implemented using PHP [11] and NuSOAP Web Services Toolkit [12]. This combination enables rapid development of such applications as Web-services. Below, an example source code of the developed service is presented:

```
$s = new soap_server;
$s -> register('sendRequest');
$s -> register('getStatus');
$s -> register('getResult');
$s -> service($HTTP_RAW_POST_DATA);
```

Besides such input parameters as the customer location, suppliers to be visited, etc., the system accepts such parameters as optimization preferences – what is more important: transportation time or costs. The user may choose his/her preference and the implemented fuzzy logic module allows finding more appropriate solution.

Presented example illustrates finding a routing plan for the same conditions but with different user preferences, namely: *minimize time*; *minimize time, then costs*; *minimize both time and costs*; *minimize costs, then time*; *minimize costs*. In (Figure 6–7) results for different choices are presented and compared. For illustration of the results a map is generated that uses the following notations. Green dots are the cities of the region. The city with red edge (Aida) is the city where the hospital is to be located – the location of the customer. The cities with blue edges are the cities where suppliers are located (Libar, Higgville, Ugwulu, Langford, Nedalla, Laki, Dado). Transportations routes are shown as lines. The grey lines are routes that are not used for transportation in the solution, the blue lines routes used for transportation, and the red lines are routes unavailable due to weather or for some other reasons. E.g., the routes through the city of Zaribe are not available because of the flooding. The colored trucks denote the routes of particular vehicles/vehicle groups.

Figure 9 represents a comparison of the routing plans created for different criteria. As it can be seen while importance of one of the parameters increases (e.g., importance for costs increases from left to right) the value of the parameter decreases (the red line with diamonds for the costs) and vice versa (the green line with squares for the time).

Conclusions

The paper describes the Web-service based approach to knowledge logistics and its applicability to transportation system management. The scalable architecture of the approach enables its extension in regard of knowledge/information sources number and, thereby, in regard of factors taken into account during complex problem solving. Implementation of the approach as a Web-service makes the system accessible from virtually any Web-enabled device (PC, PDA, or even a cell phone). Hence the knowledge base of the system can be continuously updated and the updated routing plans can be transmitted to users (e.g., truck drivers) in real time.

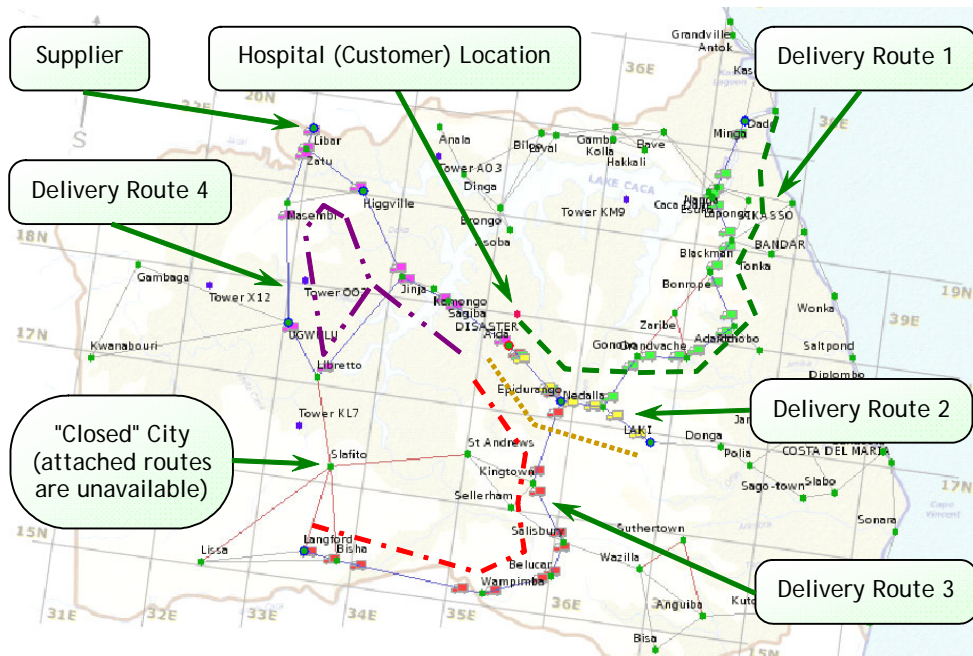


Figure 6 – Routing plan for the minimize time and minimize time, then costs preferences (in this solution four vehicles/vehicle groups are used to provide maximum of concurrency).

Acknowledgment

Some parts of the research were done as parts of the ISTC partner project # 1993P funded by Air Force Research Laboratory at Rome, NY, the project # 16.2.44 of the research program "Mathematical Modelling and Intelligent Systems", the project # 1.9 of the research program "Fundamental Basics of Information Technologies and Computer Systems" of the Russian Academy of Sciences, the grant # 02-01-00284 of the Russian Foundation for Basic Research, and the contract with Ford Motor Company. Some prototypes were developed using software granted by ILOG Inc.

References

- [1] Dey A.K., Salber D., Abowd G.D. A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications // Context-Aware Computing. - A Special Triple Issue of Human-Computer Interaction / Eds. by T.P. Moran, P. Dourish. - Lawrence-Erlbaum, 2001. Vol. 16. URL: <http://www.cc.gatech.edu/fce/ctk/pubs/HCIJ16.pdf>.

- [2] Smirnov, A., Pashkin, M., Chilov, N., Levashova, T., Haritatos, F.: Knowledge Source Network Configuration Approach to Knowledge Logistics. *Int. J. of General Systems*, Vol. 32, No. 3. Taylor & Francis Group. 2003. 251-269.
- [3] H. Baumgaertel, "Distributed Constraint Processing for Production Logistics", *IEEE Intelligent Systems*, 2000, pp. 40-48.
- [4] Smirnov, A., Pashkin, M., Chilov, N., Levashova, T.: Agent-Based Support of Mass Customization for Corporate Knowledge Management. *Eng. Applications of Artificial Intelligence*, Vol. 16, No. 4. Elsevier Science. 2003. 349-364.
- [5] ILOG Corporate Web-site, URL: <http://www.ilog.com>, 2003.
- [6] Pechoucek M., Marik V., and Barta J. CplanT: An Acquaintance Model Based Coalition Formation Multi-Agent System. *Proceedings of the Second International Workshop of Central and Eastern Europe on Multi-Agent Systems (CEEMAS'2001)*, Krakow, Poland, 2001. 209-216.
- [7] JTTP for Health Service Logistic Support in Joint Operations, 1997. http://www.dtic.mil/doctrine/jel/new_pubs/4_02_1.pdf.
- [8] DARPA Advanced Logistics Project, 2001. www.darpa.mil/iso/alp.
- [9] R.A. Rathmell, "A Coalition Force Scenario "Binni – Gateway to the Golden Bowl of Africa" (Ed.: A. Tate) *Proceedings on the International Workshop on Knowledge-Based Planning for Coalition Forces*, Edinburgh, Scotland, 1999. 115-125.
- [10] Smirnov, A., Pashkin, M., Chilov, N., Levashova, T.: Knowledge Logistics in Information Grid Environment. Zhuge, H. (ed.): The special issue "Semantic Grid and Knowledge Grid: The Next-Generation Web" of *Int. J. on Future Generation Computer Systems*, Vol. 20, No. 1. Elsevier Science. 2003. 61-79.
- [11] PHP, URL: <http://www.php.net>, 2003.
- [12] NuSOAP Web Services Toolkit. 2003. URL: <http://dietrich.ganx4.com/nusoap>.

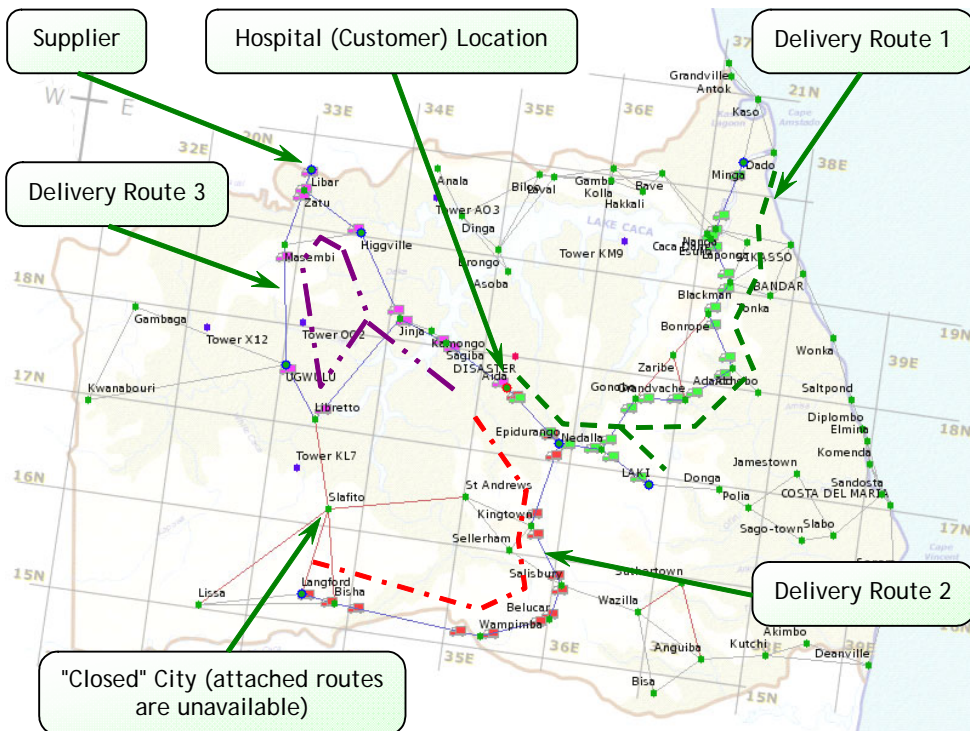


Figure 7 – Routing plan for the minimize both time and costs and minimize costs, then time preferences (in this solution three vehicles/vehicle groups are used).

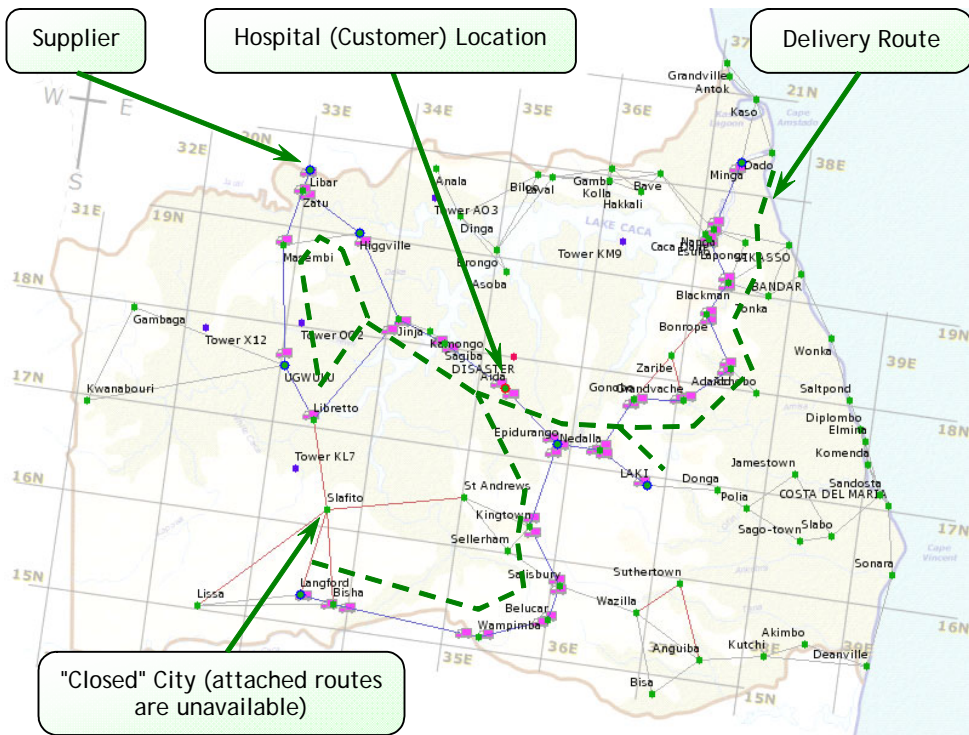


Figure 8 – Routing plan for the minimize costs preference (in this solution one vehicle/vehicle group is used to provide minimum of costs).

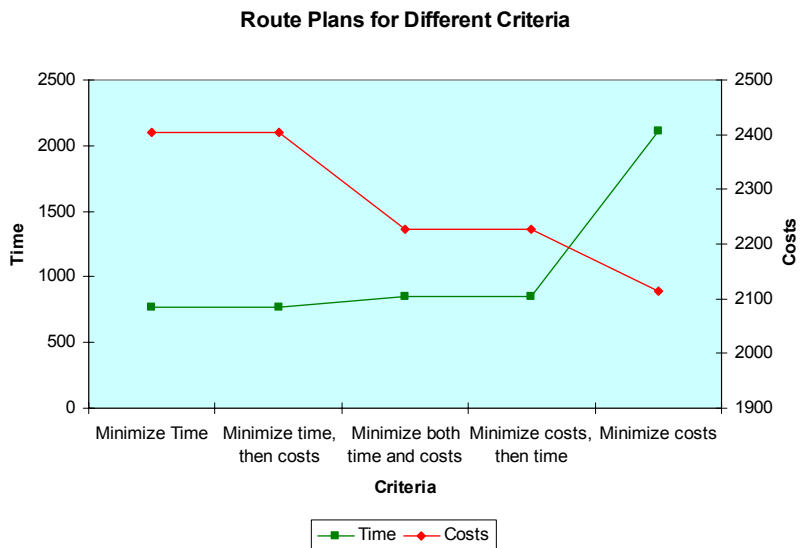


Figure 9 – Routing plans for different criteria (time and costs minimization preferences).

Biography

Alexander V. Smirnov, Prof., received his ME, his PhD and D.Sc. degrees in St.Petersburg, Russia, in 1979, 1984, and 1994 respectively. He is a Deputy Director for Research and a Head of Computer Aided Integrated Systems Laboratory at St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). He is a full professor of St.Petersburg State Polytechnical University and St.Petersburg State Electrical Engineering University. His current research interests belong to areas of corporate knowledge management, multi-agent systems, group decision support systems, virtual enterprises, and supply chain management. He has published more than 200 research papers in reviewed journals and proceedings of international conferences, books, and manuals.

Mikhail P. Pashkin received his MAM from St.Petersburg State University, Russia in 1992. He is a researcher at Computer Aided Integrated Systems Laboratory of SPIIRAS. His current research interests include multi-agent systems, ontology engineering and group decision support systems. He has published more than 40 research papers in proceedings of international conferences, books and manuals.

Nikolai Chilov received his ME in St.Petersburg State Technical University, Russia, in 1998. He is a researcher at the Computer Aided Integrated Systems Laboratory of SPIIRAS. His current research interests belong to areas of virtual enterprise configuration, supply chain management, knowledge management, ontology engineering and multi-agent systems. He is an author/co-author of more than 35 research papers published in proceedings of international conferences and books.

Tatiana Levashova received her ME degree at St.Petersburg State Electrical Engineering University in 1986. She is a leader programmer at Computer Aided Integrated Systems Laboratory of SPIIRAS. Her current research is devoted to knowledge-related problems such as knowledge representation, knowledge management, and ontology management. She has published more than 15 papers in reviewed journals and proceedings of international conferences.

Andrew Krizhanovsky received his ME in St.Petersburg State Technical University, Russia, in 2002. He is a researcher at the Computer Aided Integrated Systems Laboratory of SPIIRAS. His current research interests belong to areas of knowledge management, ontology engineering, multi-agent systems, and computer linguistics.

ЗНАНИЯ, ОСНОВАННЫЕ НА ПОНИМАНИИ, В ПРОЦЕССАХ ПРИНЯТИЯ РЕШЕНИЙ

В.А. Виттих

Институт проблем управления сложными системами РАН
443020 г. Самара, ул. Садовая, 61
vittikh@iccs.ru
тел: +7 (8462) - 32-39-27, факс: +7 (8462) - 33-27-70

Ключевые слова: *знание, основанное на понимании, онтология, неоднородный актор, координация взаимодействия акторов, централизованный, сетевой и холистический подходы к координации, социотехническая система*

Abstract

Argumentation is offered here proving the necessity of enhancing the efficiency of applying the understanding-based knowledge that reflects a joint experience and is objective and independent of an individual consciousness, in spite of the subjective nature of the knowledge acquired by autonomous heterogeneous actors. This problem solution is linked with the development of ontologies and tools for their control that facilitate the computer knowledge representation and their multiple reuse. A practical importance is emphasized of the transfer from a fragmentary approach used for the ontology construction to creation of systems of computer knowledge integration in organizations ensuring the opportunity of making a concurrent, i.e. mutually agreed, collective decisions. The advantages of a holistic approach to the actors' interactions coordination as compared to a centralized and network coordination are also shown here. The research directions are qualified related to the subject of acquisition, accumulation and application of understanding-based knowledge that is of great practical importance for the support of decision-making process in complex open systems.

Введение

Эволюционный подход к управлению сложными системами предполагает, что система не является сконструированной (заданной) априори и не имеет жесткую фиксированную структуру, а находится в постоянном развитии, в процессе которого ее целостные элементы динамически взаимодействуют между собой, устанавливая и разрывая связи под воздействием внешних или внутренних возмущающих факторов [1, 2]. В некотором исходном состоянии, когда связи между элементами отсутствуют, каждый из них функционирует автономно, и системы как таковой еще нет. Организация, формирование системы как целостности происходит при появлении *неопределенности* или проблемной ситуации, источником которой в социотехнических системах чаще всего является непредсказуемость индивидуального поведения и социальной деятельности человека [3]. Вступая во *взаимодействие*, элементы системы, включающие в себя лиц, принимающих решения, – акторов, начинают вырабатывать решения о тех действиях, которые нужно предпринять в сложившейся ситуации. Таким образом, *в процессах управления ключевую роль приобретает организация коммуникации.*

Структура коммуникации включает в себя [4]:

- участников-коммуникантов, наделенных сознанием и владеющих нормами некоторой семиотичной системы (языка);
- ситуацию, которую они стремятся осмыслить и понять;
- тексты, выражающие смысл ситуации в языке;

- цели, делающие тексты направленными, т.е. то, что побуждает субъектов обращаться друг к другу;
- процесс материальной передачи текстов.

Таким образом, *в процессе коммуникации* должен быть *понят смысл* сложившейся коммуникативно-деятельностной ситуации, поскольку *понимание предваряет* всякую, в том числе *познавательную, человеческую деятельность*. Не познание порождает потребность к пониманию, а, наоборот, потребность в понимании ведет к познанию [4, 5].

Однако при управлении открытыми сложными социальными или социотехническими системами (предприятиями, отраслями промышленности, регионами и т.п.) до использования познавательных методов, связанных с объяснением, дело чаще всего и не доходит, поскольку речь идет о принятии решений в реальном масштабе времени, когда имеется возможность только *понять ситуацию*, «ухватить ее смысл», и совершенно недостаточно времени для осуществления объяснительных процедур, предполагающих раскрытие причинно-следственных связей, построение логических умозаключений и выявление закономерностей. Иными словами, в процессах управления используются, в основном, *знания, базирующиеся на понимании*, которые в большей степени относятся к разряду интуитивных, эвристических и плохо формализуемых, а не научные знания, обладающие полнотой, логической связанностью, непротиворечивостью и представляемые обычно в виде научных теорий.

Все это находит подтверждение на практике: *«львиная» доля решений* руководителей различных рангов *принимается в результате оперативного понимания сложившейся ситуации*, а не использования традиционных научных методов, реализуемых посредством логико-методологических процедур объяснения. Поэтому ученые обвиняют управленцев-практиков в том, что они неоправданно игнорируют методы и средства, даваемые наукой, а лица, принимающие решения, видят в них малую практическую полезность, поскольку их подходы к решению задач лежат в иной плоскости – в сфере понимания, а точнее – *взаимопонимания, связанного с принятием коллегиальных решений*. А эти *проблемы – понимания и взаимопонимания в процессах принятия решений* – долгое время не привлекали внимание специалистов в области управления сложными системами. И только в последние годы можно говорить о пробудившемся интересе к ним, который выразился в появлении работ, в той или иной степени связанных с *онтологическим анализом и синтезом* при управлении сложными системами [6]. Взаимосвязь понятий «понимание» и «онтология» не случайна, поскольку философия герменевтики утверждает, что *понимание имеет онтологический статус* [5].

Иными словами, в процессах принятия решений используются два типа знаний:

- научные знания, приобретаемые с использованием объяснительных процедур;
- знания, основанные на понимании.

Последние обычно не относят к разряду научных, обладающих необходимой достоверностью. В то же самое время именно знания, основанные на понимании, чаще всего используются на практике в процессах оперативного принятия решений. Поэтому возникают методологические вопросы, связанные с противопоставлением или согласованием «понимающих» и «объясняющих» схем в познании мира человеком, онтологическим характером понимания, взаимопониманием неоднородных акторов, приобретением и использованием знаний, основанных на понимании. Данная статья посвящена рассмотрению указанных вопросов.

1 Знания, основанные на объяснении и понимании

Объяснение считается функцией познания, раскрывающей сущность одного предмета или явления (объясняемого) через другое (объясняющее), имеющего статус достоверного и «очевидного». По своей логической структуре объяснение представляет собой рассуждение или умозаключение, посылки которого содержат информацию, необходимую для обоснования такого рассуждения (умозаключения). Объяснение направлено прежде всего на установление

причинно-следственных зависимостей в изучаемом объекте с позиций гносеологии, рассматривающей процесс познания с точки зрения отношений субъекта (исследователя) к объекту или в категориальной оппозиции «субъект-объект». В европейской культуре функция объяснения постепенно закрепила за *научным знанием* [4], которое по этой причине можно рассматривать как *знание, основанное на объяснении*.

Понимание, как процедура постижения смысла, не вписывается в субъект-объектную познавательную схему [4]. Феномен понимания не опирается на классическую гносеологию с ее внеположенностью субъекта объекту и использованием метода объяснения, основным содержанием которого является подведение индивидуального под всеобщее, то есть поиск закономерностей [5].

Однако прояснение смысла бытия, по М. Хайдеггеру, возможно только путем осознания смысла самого вопрошающего, поскольку бытие становится доступным только через человеческое присутствие. *Бытие «Я» как «бытие в мире» становится принципиально коммуникативным*, т.е. индивидуальное бытие содержит бытие другого [4]. Иными словами, если для классической философии типична характеристика индивидуального сознания через его направленность на объект, то для второй половины XX века центральной становится его характеристика через стремление к другому, т.е. *субъект-объектные отношения заменяются субъект-субъектными*.

Один человек не в состоянии понять мир, т.е. осознать многообразие смыслов и значений культурно-исторического мира, которому он принадлежит. Поскольку он может понять его лишь частично, человек воспринимает от других то, чего не достает в его собственном опыте. На этом основании Э. Гуссерль ввел понятие интерсубъективности как структуры индивидуального сознания, отвечающей факту существования других индивидов [5], или особой общности между познающими субъектами, условия взаимодействия и *передачи знания одного для другого* [4].

Таким образом, *«сущность смысла* открывается в его коммуникативной природе. *Гарантом его «объективности» является не тот факт, что он адекватно отображает реальный мир, а то, что он является выражением коллективного опыта, передаваемого культурой и существующего в этом плане независимо от индивидуального сознания»* [5].

В отличие от традиционной трактовки понимания как процедуры обнаружения смысла текста в процессе его интерпретации, реконструирующей изначальный его замысел, во второй половине XX века *понимание* стало рассматриваться как *эвристическая познавательная процедура, дающая приращение знания*, а не только восстанавливающая изначальное [4]. Поэтому, наряду со знаниями, базирующимися на объяснении, можно говорить о *знаниях, основанных на понимании*.

Следует отметить, что на Востоке (например, в Китае) знание также не преследовало цель объяснить мир. Целью было понимание мира, что, кстати сказать, не способствовало развитию точных наук, как это было на Западе. Китайское философское мышление всегда характеризовалось отчетливо выраженным практицизмом, поскольку знание предмета включало не только его описание, но и предписание к действию: обладать знанием значит прежде всего знать свое дело.

2 Онтологический статус понимания

Мы говорим о понимании коммуникативно-деятельностной ситуации, в которую «погружен» человек, полагая, что он не только изучает ее «изнутри» и взаимодействует с другими людьми, но и оказывает воздействие на нее [2]. Эта сращенность человека с бытием и придает пониманию свойство естественно складывающегося процесса, предполагающего прежде всего «онтологическое (а не гносеологическое) самоопределение субъекта, будучи проблемой практического отношения и практического разума» [4]. Прежде чем объяснить мир, человек пости-

гает смысл происходящего «интегрированно», воспринимая и описывая мир как целостность, таким, какой он есть, не прибегая к использованию инструментов в виде разнообразных формальных теорий, которые сужают и дробят на отдельные составляющие спектр его представлений о мире. Иными словами, *понимание органично связано с понятием «онтология»*, которая трактуется как «описание и организация того, что есть, множество вещей, которые существуют, и которая определяет, как эти вещи взаимосвязаны» [7].

Онтологический статус понимания был определен М. Хайдеггером, который представлял понимание как специфическое отношение к действительности, в котором человек выступает толкующим себя бытием (т.е. «понимающим бытием»). Такая трактовка понимания закрепились затем в герменевтике Г. Гадамера. Герменевтика становится у Г. Гадамера онтологией, основанием которой является язык. Полагая язык в качестве среды герменевтического опыта, Г. Гадамер исходил из того, что «языковым (и потому понятным) является сам человеческий опыт мира. Сам мир выражает себя в языке» [4].

Онтологический характер понимания, с которым мы сталкиваемся как в теоретических исследованиях, так и в повседневной жизни, *предваряет всякую человеческую деятельность* в качестве предпонимания [5]. Понятие «*предпонимание*» введено М. Хайдеггером, который выразил в нем *развертывание понимания* как онтологического определения человеческого бытия [4]. Любой ситуации, в которой оказался человек, предшествует некоторое предпонимание, выраженное в форме онтологий и представляющее собой знания, на базе которых осуществляется процедура постижения смысла сложившейся ситуации. На основе этого понимания приобретаются новые знания, пополняющие онтологии, которые становятся предпониманием к следующей ситуации и т.д.

Знания, основанные на понимании и представляемые в форме онтологий, даны человеку, как члену социума, изначально, априори, определяя ему смысловой порядок, который кажется человеку естественным видением мира. Поэтому *каждый объект* из своего окружения *человек может наделить смыслом*, в результате чего формируется множество онтологий – описаний объектов и отношений между ними. Понимание объекта или ситуации не статично, а изменяется со временем, поэтому трансформируются и *онтологии, в которых отражается*, таким образом, *текущее состояние осмысления действительности человеком*. И поскольку в процессах принятия решений понимание происходит в условиях коммуникации акторов, на передний план выдвигается проблема достижения взаимопонимания между ними.

3 Взаимопонимание неоднородных акторов

Взаимопонимание обычно трактуется как согласие, общее для всех сторон отношение к кому-либо или чему-либо. Если акторы однородны, т.е. их поведение и предпочтения идентичны в любых ситуациях, то взаимопонимание заложено в них изначально. Однако *гипотеза об однородности акторов является далекой от реальности идеализацией*, а применительно к экономической теории она «столь слаба, что лишь немногие экономисты до сих пор отстаивают ее» [8], поскольку в ее основе «лежит идея о метафизическом единстве человеческих желаний» [9]. Заслуживает критики и централизованное планирование, потому что «условием его эффективности является стационарность мира, состоящего из однородных экономических агентов» [8].

Взаимопонимание неоднородных акторов, имеющих различные интересы и предпочтения, является неотъемлемой характеристикой процесса децентрализованного принятия решений. Здесь уже взаимопонимание *может рассматриваться не только как согласие или единомыслие* (understanding), *но и как своеобразное «единство противоположностей»*, обозначаемое в английском языке как “reciprocal understanding”: если один актер (экономический агент) имеет возможность удовлетворить потребность другого актора, то договоренность об оказании соот-

ветствующей услуги можно интерпретировать как достижение взаимопонимания между ними [10]. Таким образом, неоднородные акторы могут находить взаимопонимание в двух смыслах:

- согласия как общности позиций (точек зрения);
- единства как «притяжения различных» (а не одинаковости).

При этом для реализации процесса достижения взаимопонимания могут использоваться *три способа координации взаимодействия акторов – централизованный, сетевой и холистический*. *Централизованный* подход, связанный с жесткой иерархией и административным командованием, исходит из того, что координация осуществляется по принципу *примата коллективных интересов* по отношению к индивидуальным: правила поведения авторов навязываются им «сверху». *Сетевая организация*, наоборот, предполагает установление двусторонних связей по обоюдному согласию акторов *для достижения их индивидуальных целей*. Каждый актор кооперируется с другими акторами исключительно из соображений своей собственной выгоды. Коллективные интересы в расчет не принимаются. Взаимопонимание между партнерами достигается либо в случае совпадения их целей (интересов), либо тогда, когда один из них имеет возможность удовлетворить потребность другого. На основе такого взаимопонимания затем устанавливаются двусторонние связи. При потере взаимопонимания связи разрываются.

Холистический подход к координации, занимающий промежуточное положение между централизованным и сетевым, исходит из допущения, что *связь устанавливается* не между двумя акторами, а *между актором и коллективом* (группой) акторов. Его приверженцы усматривают в механизмах решений, принимаемых на индивидуальном уровне, прообраз коллективных форм, с которыми связан индивид (актор). Поэтому основной проблемой, связанной с использованием холистического подхода, является определение характера отношений между актором и коллективом акторов [8]. Например, в [9] подчеркивается, что задача максимизации национального богатства должна быть воспринята на уровне каждого актора и найти отражение в индивидуальных функциях предпочтений.

Холистические идеи в экономике определяют условия эффективного использования права собственности, которые заключаются в отказе от части этого права в пользу коллектива, представленного государством или муниципальными службами. К примеру, чтобы иметь право использовать свой собственный автомобиль, необходимо согласиться с ограничением скорости, мест парковки, разнообразных технических условий и даже ограничением дней, когда его можно использовать [8].

Иными словами, *достижение взаимопонимания* между неоднородными акторами при холистическом подходе *осуществляется путем соглашения о правилах кооперации*, снижающих неопределенность потенциальной конфронтации между акторами до такой степени, *что становится возможным действовать в интересах отдельного актора*, направленных на удовлетворение его субъективных нужд. Тем не менее введение правил кооперации не исключает полностью риска конфронтации, поскольку кооперация не поглощает полностью неопределенность. Правила могут лишь снизить неопределенность до некоторого приемлемого уровня в каждой конкретной ситуации и на определенном промежутке времени [3].

Следует отметить, что холистический способ координации взаимодействия акторов наиболее распространен на практике, поскольку централизованный и сетевой подходы «в чистом виде» не используются. Поэтому в реальной жизни приходится постоянно решать достаточно сложную проблему нахождения компромисса между индивидуальными и коллективными интересами акторов.

4 Приобретение и использование знаний, основанных на понимании

Знания, основанные на понимании, играют ключевую роль в науках об искусственном – технических, экономических, социальных и т.п. Специфика таких инженерных знаний состоит

в том, что в отличие от естественнонаучных знаний, «извлекаемых» исследователями из объектов по классической субъект-объектной схеме, они приобретаются исследователями-инженерами (актерами), выполняющими познавательную-созидательную функцию, с использованием субъект-субъектных отношений [11]. Актеры неодинаковы, не вписываются в какую-то жесткую однородную структуру и не имеют единообразного понимания сложившейся ситуации, т.е. в широком смысле являются неоднородными. Тогда понятие «реальный мир» включает в себя прежде всего *сумму представлений акторов о самих себе и о других*, об их собственной позиции, о мире, в котором они действуют [12]. Поэтому «децентрализованная экономика имеет субъективистскую природу: она формируется в результате взаимодействия различных субъективных начал» [8].

Многообразие знаний неоднородных акторов образует «рассеянное знание» [13]; причем само понятие «знание» Ф. Хайек трактовал как «способ существования развивающейся мысли, которая при столкновении с новыми обстоятельствами позволяет отдельному инженеру быстро находить новые решения» [14]. В своих фундаментальных работах Ф. Хайек (еще в середине прошлого века!) рассматривал *рынок как систему распространения знания*, позволяющую использовать такой объем информации, получение и переработка которого невозможны в рамках системы централизованного планирования. Исходя из того, что *знание акторов постоянно меняется и преобразуется*, Ф. Хайек различал *собственно знание и процесс его приобретения* [8].

Развитие общества Ф. Хайек представлял как процесс эволюции и взаимодействия «спонтанных порядков» - социальных институтов, моральных традиций и практик, складывающихся на основе принципов самоорганизации, полагая, что *координация деятельности акторов* в рамках «спонтанных порядков» осуществляется путем соблюдения универсальных *правил поведения с одновременным предоставлением каждому актору определенной автономии* [4]. Таким образом, он постулировал холистический подход к координации.

Опираясь на изложенную ретроспективу можно сделать вывод о том, что в упомянутых работах термин «знание» обозначает знание, основанное на понимании, которое исходно приобретает автономными неоднородными актерами (и в этом смысле знание субъективно), но в результате взаимодействия акторов с помощью некоторой «системы распространения знания» (например, рынка) рождается новое знание, выражающее уже коллективный опыт и существующее в этом плане объективно, независимо от индивидуального сознания. И если обратиться к реальной действительности, то следует признать, что именно такая модель работы со знаниями, основанными на понимании, реализуется на практике. Однако *эффективность приобретения и использования таких знаний остается чрезвычайно низкой*, поскольку эти процедуры осуществляются интуитивно, эвристически и плохо формализуемы.

Важным шагом на пути решения этой проблемы явилась разработка *онтологий и средств управления ими* [15, 16], которые позволяют представлять знания в компьютерах и многократно их применять. Однако, этого оказалось недостаточно, поскольку при этом необходимо перейти от фрагментарного построения онтологий к созданию в организациях систем компьютерной интеграции знаний, обеспечивающих возможность согласованного, основанного на достижении взаимопонимания, принятия коллегиальных решений [17]. В организациях должны начать работать инженеры по знаниям, разрабатывающие онтологии, относящиеся к отдельным актерам, и онтологии, отражающие коллективные интересы групп неоднородных акторов. Должен быть разработан механизм приобретения знаний, основанных на понимании, учитывающий их динамику, а также созданы средства доступа к накопленным знаниям, обеспечивающие возможность их применения различными категориями пользователей в процессах принятия решений.

Сведение с помощью информационно-коммуникационных технологий в единую развивающуюся систему знаний, основанных на понимании, позволяет интерпретировать ее как некоторую инженерную теорию [18, 19], отличающуюся от классической теории, о которой обычно судят как о концептуальной системе, обладающей внутренней непротиворечивостью и

удовлетворяющей некоторой совокупности исходных аксиом. Однако, если *теория адресуется к открытым системам*, состоящим из неоднородных элементов, необходимо строить ее на базе *онтологических представлений*, т.е. нужно «обратиться к *принципу описательности*, отказавшись от чисто аксиоматического принципа построения теории, который в действительности нацелен лишь на то, чтобы оградить тех, кто ему следует, от волнующих встреч с реальным миром» [8]. С этой точки зрения изложенные принципы могут быть применены при создании теории управления открытыми сложными социотехническими и экономическими неоднородными системами.

Заключение

Повышение эффективности приобретения и использования знаний, основанных на понимании, в процессах принятия решений является одной из наиболее актуальных проблем управления сложными системами. Эти знания выражают коллективный опыт и существуют этом плане объективно, независимо от индивидуального сознания, несмотря на то, что они, будучи приобретаемыми автономными неоднородными акторами, имеют субъективное происхождение.

Для координации взаимодействия акторов могут использоваться три подхода – централизованный, сетевой и холистический. Централизованный подход, связанный с жесткой иерархией и административным командованием, исходит из того, что координация осуществляется по принципу примата коллективных интересов по отношению к индивидуальным. Сетевой способ предусматривает установление двусторонних связей акторов на основе их собственной выгоды, а коллективные интересы в расчет не принимаются. Холистический подход к координации имеет преимущество перед централизованным и сетевым, поскольку исходит из допущения, что связь устанавливается не между двумя акторами, а между актором и коллективом (группой) акторов: в механизмах решений, принимаемых на индивидуальном уровне, усматривается прообраз коллективных форм, с которыми связан индивид (актор). При этом основной проблемой, связанной с использованием холистического подхода, является нахождение компромисса между индивидуальными и коллективными интересами акторов.

Решение проблемы повышения эффективности приобретения и использования знаний, основанных на понимании, предполагает разработку онтологий и средств управления ими, которые позволяют представить знания в компьютерах и многократно их применять. Важно иметь в виду, что при этом должен быть осуществлен переход от фрагментарного подхода к построению онтологий к созданию в организациях систем компьютерной интеграции знаний, обеспечивающих возможность согласованного, основанного на достижении взаимопонимания, принятия коллегиальных решений.

Сведение с помощью информационно-коммуникационных технологий в единую развивающуюся систему знаний, основанных на понимании, открывает новые возможности создания теории управления сложными социотехническими и экономическими неоднородными системами, которая должна дать ответы на следующие вопросы:

- Какими способами должны приобретаться знания, основанные на понимании?
- Каковы методы оценки и повышения степени их достоверности?
- Каким образом должна быть организована субъект-субъектная процедура приобретения и извлечения знаний, основанных на понимании, из акторов и из баз данных?
- Каковы способы представления в компьютерах знаний, основанных на понимании?
- С помощью каких программных средств эти знания могут быть эффективно использованы в процессах принятия коллегиальных решений?
- Каковы методы и средства построения эволюционирующих баз знаний, учитывающих динамику знаний, основанных на понимании, в сложных открытых системах?

Получение ответов на эти и другие вопросы, относящиеся к проблематике приобретения, накопления и использования знаний, основанных на понимании, является предметом научных исследований, имеющих важное практическое значение для поддержки процессов принятия решений в сложных открытых системах.

Список литературы

- [1] Виттих В.А. Эволюционное управление сложными системами. – Известия Самарского научного центра РАН, том 2, №1, 2000, с.53-65.
- [2] Виттих В.А. Целостность сложных систем. – Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции. Самара, Самарский научный центр РАН, 2002, с.48-58.
- [3] Küppers G. Self-organization – The Emergence of Order. From local interactions to global structures. – <http://www.uni-bielefeld.de/iwt/sein/paper no2, pdf, July 1999/>
- [4] Всемирная энциклопедия. Философия. – М.: АСТ, Мн.: Харвест, Современный литератор, 2001.
- [5] Современный философский словарь. – Москва-Минск, «ПАНПРИНТ», 1998.
- [6] Виттих В.А. Онтологический анализ и синтез при управлении сложными открытыми системами. – Проблемы управления и моделирования в сложных системах: Труды V Международной конференции. Самара, Самарский научный центр РАН, 2003, с.50-60.
- [7] Merrill G.H. The Babylon Project: Towards an Extensible Text-Mining Platform. – IEEE IT Pro, March-April 2003, IEEE Computer Society, 2003.
- [8] Сапир Ж. К экономической теории неоднородных систем (опыт исследования децентрализованной экономики). – М., Государственный университет – Высшая школа экономики, 2001.
- [9] Myrdal G. The Political Element in the Development of Economic Theory. – Cambridge Mass.: Harvard University Press, 1954.
- [10] Виттих В.А., Скобелев П.О. Мультиагентные модели взаимодействия для построения сетей потребностей и возможностей в открытых системах. – Автоматика и телемеханика, №1, 2003, с.177-185.
- [11] Виттих В.А. Инженерная эпистемология. – Проблемы управления и моделирования в сложных системах: Труды III Международной конференции, Самара, Самарский научный центр РАН, 2001, с.92-101.
- [12] Mäki U. How to combine Rethoric and Realism in the Methodology of Economics. – Economics and Philosophy, v.4, 1988, p. 353-373.
- [13] Hayek F. The use of knowledge in society. – American Economic Review, vol. XXXV, №4, 1945, p.519-530.
- [14] Hayek F. Individualization and Economic Order. – Chicago, Mich.: University of Chicago Press, 1948. (русский перевод: Хайек Ф.А. Индивидуализм и экономический порядок. – М.: Изограф.: Начала-Фонд, 2000).
- [15] Смирнов С.В. Онтологический анализ в системах компьютерного моделирования. – Проблемы управления и моделирования в сложных системах: Труды V Международной конференции, Самара, Самарский научный центр РАН, 2003, с.102-107.
- [16] Левашова Т.В., Пашкин М.П., Смирнов А.В., Шилов Н.Г. Управление онтологиями (базами знаний). – Известия академии наук. Теория и системы управления, №4, 2003, с.132-146.
- [17] Виттих В.А. Управление открытыми системами на основе интеграции знаний. – Автометрия, №3,1998, с.38-49.
- [18] Vittikh V.A. Engineering theories as a basis for integrating deep engineering knowledge. – Artificial intelligence in Engineering, vol.11, №1, 1997, p.25-30/
- [19] Виттих В.А. Интеграция знаний при исследованиях сложных систем. – Известия академии наук. Теория и системы управления, №5, 1998, с.132-139.

ANALYSIS OF ACTIVE RESPONSE IN THE IMMUNE SYSTEM WITH COMPUTER NETWORK CONSIDERATIONS

V.A. Skormin, D.H. Summerville, O.I. Shirayeva, J.S. Moronski

Binghamton University, Binghamton NY 13802, USA

vskormin@binghamton.edu

tel: (607)777-4013, fax: (607)777-4464

Key words: immune system, information attack, active response, computer network

Abstract

Реакция иммунной системы организма на вторжение чужеродных белков рассматривается в качестве положительного примера работы механизма защиты сложной системы от информационной атаки. Описаны специфические составляющие этого процесса. Они послужили шаблоном для создания математической модели. В целом этот механизм рассматривается как циклическое, нелинейное, отрицательное замыкание с обратной связью. Выявлены основные условия двух крайних результатов такого нападения – смерти либо полного восстановления. Описано правило, по которому неизбежно полное восстановление. Представлена хронология недавней информационной атаки на международное компьютерное сообщество, и установлено сходство отдельных стадий нападения на компьютерную сеть и на иммунную систему. Основной вывод состоит в том, что управляющее правило, обеспечивающее обороноспособность иммунной системы, может быть с успехом использовано для отражения нападений на компьютеры.

Introduction

Modern immunology provides a detailed but rather qualitative description of the active response of the immune system to any entity invading the organism and recognized as an antigen by the immune cells. In a similar way that a malicious computer program is constructed from the same instructions as legitimate software but sequenced in a fashion that makes it malicious, an antigen is constructed from the same building blocks as the cells of the host but differently sequenced. In many ways, the effects of an antigen on a biological organism are similar to those caused by some information attacks on computer networks. This is the major reason for considering the active response of the immune system, honed to perfection by million-year evolution, as an ideal mechanism for protecting computer networks from information attacks.

The specific immune response is the main mechanism enabling the immune system to destroy cells of the intruding antigen. The outcome of this process (recovery, chronic infection or lethality) greatly depends on the time between the infection and detection of the antigen, and growth rates of the antigen cells and immune cells competing for the limited resources of the host. After the antigen has been defeated, the residual concentration of specialized fighter cells slowly decreases providing high immunity to similar infection. We attempt to describe the active response mechanism of the immune system in terms of a negative-feedback closed-loop system and establish nonlinear differential equations of its particular components thus resulting in a mathematical model of the immune response. Subjected to analytical techniques offered by modern control theory, this model will enable us to establish the conditions for three possible outcomes of such interaction, full recovery, chronic infection, and lethality, and to formulate a control law assuring the full recovery outcome.

We will demonstrate that an equivalent of every stage of the immune response could be observed in our experience with computer epidemics. Consequently, the model of the immune response, describing the principle of operation of a successful defense mechanism against information attacks, has

great potential for the analysis and synthesis of defenses for a computer network. The control law synthesized using the methodology of advanced control theory to assure a full recovery outcome for the immune system can be reformulated in terms of the characteristics of computer networks and interpreted as a set of instructions to a network manager.

While computer networks became one of the critical and increasingly vulnerable components of the national infrastructure, our research addresses this reality presenting a methodology for the analysis and synthesis of the information security systems for computer networks that emulates the genetically optimized mechanism of active immune response. It is said that the immune response is genetically optimized for its biological environment; although this is different from a computer network, this difference narrows with every recent advancement in nano-technology and synthetic protein-type computing media that will result in large-scale integrated computer networks whose operation will resemble the living tissue. The presented effort presents an innovative approach to mathematical modeling, computer simulation, analysis, design and control of complex systems resulting from these emerging technologies, and provides the visualization and important insight into the interrelation of physical phenomena behind their operation.

This work is a logical continuation of the research conducted by the authors under the Air Force funding. In 2001 they engaged in the project BASIS (Biological Approach to System Information Security) aimed at the analysis of the aspects of the specific immune response that have the potential for the implementation in the new generation of computer network security systems [1, 2, 3]. This project has resulted in the “engineering view” of the immune mechanisms and prompted several new concepts in computer security. First, utilizing the modern findings of immunology the concepts such as formal protein and formal immune network were formulated presenting a novel approach to modeling the immune system. The capability of the immune system to make a distinction between the “self” from “non-self” with high degree of dependability by determining the binding energy between proteins was emulated using the methodology developed in matrix analysis. The computational approach involving the concept of binding energy was proposed as *Immunocomputing* for the implementation in novel computing media and application for the solution of a wide class of computationally intensive problems [4].

The second BASIS-inspired concept was the detection of malicious codes by detecting its “gene of self-replication”. Indeed, a high percentage of information attacks are perpetrated by deploying computer viruses and worms, which result in very costly and destructive “epidemics”. Spread of malicious codes is achieved by the built-in ability to self-replicate through the Internet and computer media. Since most legitimate codes do not self-replicate, and the number of ways to achieve self-replication is limited to the order of fifty, the detection of malicious codes could be reduced to the detection of the “gene of self-replication” in the code in question. This research effort is on the way and, according to current results, is highly successful [5].

1 Anatomy of the active immune response

The specific immune response is the main mechanism enabling the immune system to destroy cells of the intruding antigen. This is accomplished by multiplying, on demand, fighter cells that are uniquely equipped for counteracting this particular antigen by carrying its genetic sample. The immune system is prepared to counteract practically any antigen as it contains cells that specialize in at least 10^{15} various genotypes. However, the actual ability of the immune system to destroy an intruder depends on its ability to generate specialized fighter cells at the necessary rate, i.e. to actively respond to the intruder.

Active response of the immune system includes several stages. It starts from the intrusion of the antigen cells in the biological organism. Intruding cells quickly multiply and their concentration exponentially increases thus increases the probability of a physical contact of an antigen cell with a specialized immune cell capable of recognizing it as an antigen. The initial concentration of these specialized

cells depends on the previous exposures of the organism to this antigen (i.e. acquired immunity). The detection of the antigen triggers the process of exponential proliferation of immune cell-fighters specialized to destroy the antigen cells. Multiplying antigen cells and multiplying immune fighter cells compete for limited resources of the biological organism. It could be seen that the outcome of this process (recovery, chronic infection or lethality) greatly depends on the time period between the moment of infection and the moment of detection of the antigen. When the proliferation of the antigen cells goes too long before detection, its cells consume a greater share of the resources of the host thus preventing the fighter cells from sufficiently multiplying, leading to lethality. A high initial concentration of specialized fighter cells (after the organism has been immunized for a particular infection) facilitates early detection of the antigen and prevents it from overwhelming the immune defenses. After the antigen has been defeated, the residual concentration of specialized fighter cells slowly decreases providing high immunity to similar infection. It could be seen that at certain conditions, parity between proliferating antigen and fighter cell could be achieved leading to chronic infection.

2 Mathematical model of the immune response

The mathematical model of immune response reflects the basic concepts and phenomena of immunology and describes the dynamics of the immune process at the organism level. The individual components of a living organism's immune response are not essential to the analysis of the dynamics of the response to an antigen attack, therefore only the basic mechanisms of the protective reaction of an organism without distinction between cellular and humoral responses will be reflected. Included in the generalized definition of protective cells are the components of cellular/lymphoid systems as well as the humoral/immunoglobulin system. This includes leukocytes, lymphocytes, antibodies/immunoglobulin and cellular structures that are capable of neutralizing a given antigen. Antigen is defined as any organism or material alien to a targeted system that is capable of causing an immune response. Based upon these preliminary assumptions the mathematical model of the immune response presents a combination of the following components.

Limited resources of the organism: Proliferation of the antigen cells as well as the immune cells in a living organism is possible only if there are available resources. Antigen and immune cells compete for these resources while consuming them. These resources are being restored by various processes within the organism and are maintained at some level that slowly decreases due to ageing. Mathematical model of the immune response will include a non-negative variable providing a quantitative representation of the vital resources of an organism. A single-loop control mechanism responsible for maintaining the level of resources and special parameters regulating the rate of ageing will be introduced in the model.

Initial concentration of specialized immune cells: The concentration of the specialized immune cells prior to attack will be emulated by a constant, in the case of "normal" low immunity, or slowly decreasing exponential, in the case of high immunity acquired after previous infection by the same antigen.

Proliferation of the antigen cells: The initial concentration of the antigen cells will be represented by some amount in the moment of infection representing the severity of infection. Antigen concentration dynamics after the moment of infection corresponds to models of population growth. One of the elementary models of a population growth was introduced by T. Malthus based upon the tendency of a population to increase in a geometrical progression. In nature, many life forms are capable of propagating in a geometrical progression, however, factors such as competition for resources, disease, and death (natural and forced) present obstacles for the sustaining growth. Verhulst first introduced the mathematical form of an S-shaped growth curve also referred to as the Verhulst logistical curve. The inclinations of S-curve grow exponentially in the beginning then gradually flatten towards zero. At large time values the curve converges to a horizontal line represented by which describes the equilibrium value of a population size.

The equation, which describes the rate of change in concentration of an antigen after the moment of infection, could be developed using one of the equations of natural dynamics [10]. In the case of antigen introduction, a cascade increase in the concentration of antigen takes place due to absorption of vital resources of the host. In the beginning, the amount of available resources does not influence the rate of increase of antigen until the concentration of the antigen exceeds some threshold. Upon reaching this level, the resources of the host organism will not be sufficient to sustain the cascade growth of the antigen, therefore, the growth rate of the antigen will gradually decrease to a minimum level. The given process can be described by a nonlinear logic function that reflects only the presence or absence of resources sustaining the increase of the antigen concentration.

Detection of the antigen: Detection of the antigen is the event visualized as the first physical contact of the antigen cell and the specialized immune cell “equipped” to recognize this antigen. It could be seen that the probability of this event increases with the increase of the concentrations of both the antigen cells and the specialized immune cells. Prior to the moment of infection, the concentration of the specialized immune cells could be viewed as a slow decreasing exponential function, practically a constant. It is the multiplication of the antigen cells that causes the increase of the probability of the antigen detection with time making this event inevitable over some period of time. One could assume that the *detection time* can be defined as the time period during which the probability of detection reaches some sufficiently high value. It could be seen that in the case of elevated immunity of the organism, i.e. high initial concentration of the specialized immune cells after previous attack by the same antigen, the detection time is shorter than in the case when the organism has never been exposed to the antigen. Shorter detection time also causes immune defenses to be activated at an earlier stage of the antigen multiplication process that significantly increases the chances of the multiplying immune cells to successfully compete for the limited resources of the organism.

Interaction between the antigen and immune cells: From the point of infection, antigen cells are destroyed by the specialized immune cells that results in the reduction of the antigen cell concentration. Consequently, the differential equation of the concentration of the antigen cells would reflect the following phenomena,

- a) Multiplication of the antigen cells affected by the available resources of the organism;
- b) Resources of the organism affected by the proliferating antigen and immune cells;
- c) Neutralization of the antigen cells by immune cells dependent on the concentrations of both cells, the probability of a binding between an antigen and an immune cells, and a number of immune cells participating in the neutralization of one antigen cell [11].

The probability of a binding also depends on the concentrations of the antigen and immune cells and the resources of the organism [12].

Dynamics of the concentration of the immune cells: The concentration increase of the immune cells is triggered by the detection of the antigen and takes place only when there is a surplus of resources. The rate of increase is also dependent on the availability of resources. This reality is well known in immunology: low initial immunity level resulting in long detection time and/or large amount of the antigen introduced at the moment of infection allow the antigen cells to consume a large share of the resources of the host thus preventing the immune cells from proliferating at the necessary rate.

When developing the equation that describes the dynamics of concentration of the immune cells in the presence of antigen, it is necessary to take into account that the organism concentrates its response primarily in the region where most of the antigen cells are located. With a small antigen concentration, there is a weak stimulation of lymphocytes. If the antigen concentration is large, many lymphocytes may reach the end of their ability to propagate [12].

The differential equation of the concentration of the immune cells reflect,

- The cascade growth of the concentration of immune cells binding with an antigen stimulated by the presence of an antigen;
- The decrease of the concentration of immune cells as a result of interaction with antigen;
- The dynamics of the available resource of the organism.

Finally, *the mathematical model of the immune response system* can be defined as a system of nonlinear differential equations interrelating three time-dependent *variables* that represent

- Vital resources of the host organism,
- Concentration of the antigen cells
- Concentration of the specialized immune cells.

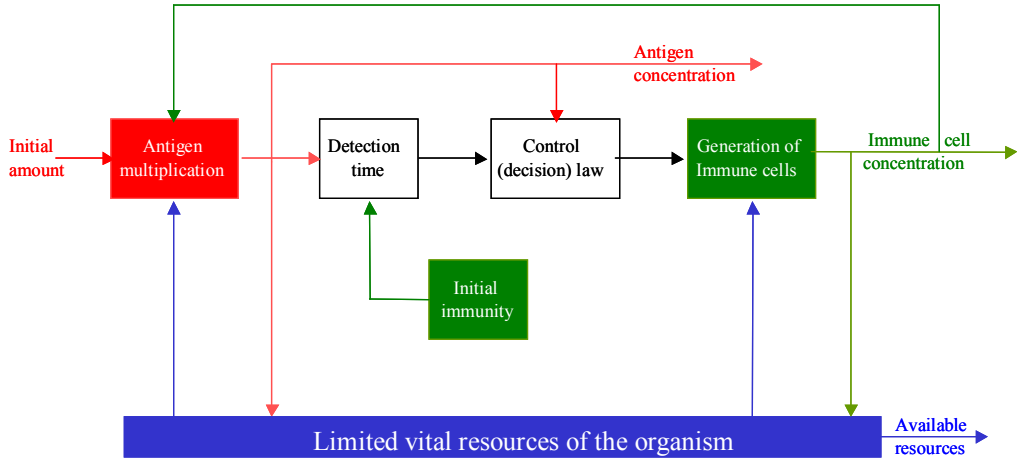


Figure 1 – Simulation model of the immune response.

The incomplete list of *parameters and constants* of the mathematical model includes

- rate of natural ageing of the organism (resource reduction),
- coefficients of an organism's recovery rate,
- amount of resources of the organism per unit of antigen cell concentration,
- amount of resources of the organism per unit per unit of immune cell concentration,
- rate of cascade growth of the antigen concentration due to availability of resources,
- rate of the concentration change of the antigen due to unavailability of resources,
- reduction of the antigen concentration due to the interaction with immune cells,
- rate of growth of the immune cell concentration under the availability of resources,
- rate of decrease of the immune cell concentration when resources are unavailable,
- reduction of the immune cells concentration due to the interaction with the antigen.

A preliminary version of such a model established under the Air Force funding [1] and implemented in simulation environment providing the means for the quantified visualization of particular components of the immune response. It is understood that the model is lacking some details and the choice of its parameters is not fully supported by literature. Nevertheless, the model allows for simulation analysis of the conditions leading to three possible outcomes of the antigen attack on the living organism.

3 Outcomes of the immune response

The immune response is modeled as a continuous dynamic process [1]. The biological variability or casual statistical fluctuations was disregarded thus the model represented average values of the ap-

appropriate variables. The qualitative side of the particular phenomena was reflected by the choice of coefficients (gains).

3.1 Primary and secondary immune response with the full recovery outcome

This outcome results in the elimination of the entire population of the antigen cells and is expected when the organism's resources are sufficient for supporting the necessary rate of multiplication of the immune cells. After the initial infection, the antigen cells multiply undetected by the immune system. This period is referred to as the incubation period of the disease. The length of this period can vary as antigen may lie dormant waiting for some “trigger event” to awaken it. Once detected, there is an additional period of time during which the immune cells and antigen multiply without significant interaction: at the beginning stages of the antigen attack the immune system concentrates its efforts on creating appropriate cells to respond to the newly discovered infection. When the concentration of antigen reaches some level, the second phase of the immune response, the growth phase, begins with the occurrence of the antibodies in the blood whose concentration increases over a several day period to the highest possible level. This is the most active phase of the immune response to an antigen. At this phase, the concentration of antigen cells begins to decrease as a result of the annihilating by immune cells. In addition, the growth of the concentration of immune cells slows the rate of multiplication of the antigen cells that ultimately results in the elimination of antigen.

After the population of antigen is nearly suppressed, the final phase of the immune response begins. This phase is referred to as the attenuation of the immune response. It ends with the residual level of specialized immune cells being much higher than prior to the antigen attack resulting in the high (acquired) immunity of the organism to this particular antigen.

During all three phases of the primary immune response, populations of antigen and protective cells are consuming resources of the organism causing their level to decrease, especially during the period of the greatest concentration of antigen and immune cells. It is important that the resource level does not become the factor limiting the multiplication rate of the immune cells that becomes the condition for the full recovery outcome.

The secondary immune response occurs with a repeated infection of an organism by the same antigen. Then due to high initial concentration of the specialized immune cells (high immunity) the antigen is detected virtually immediately and all stages of the immune response take place very fast with very little impact on the resources of the organism.

The simulation analysis results illustrating the full recovery outcome of the immune response are shown below.

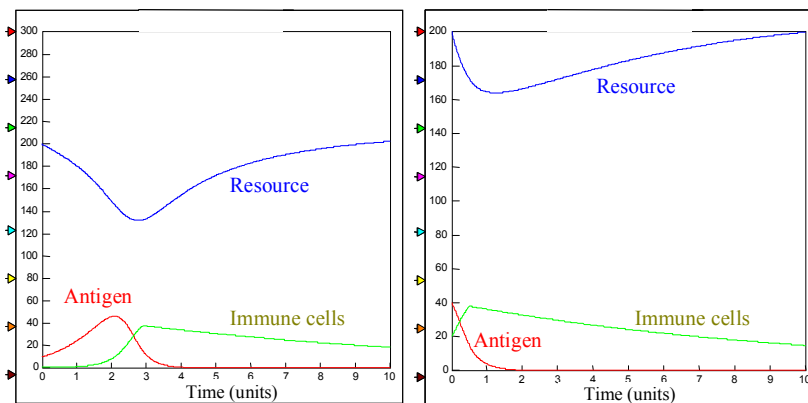


Figure 2 – Primary and secondary immune responses with full recovery.

3.2 Failure of the immune response: lethality or chronic illness

A severe infection in combination with low initial immunity and insufficient immune response results in lethality or chronic illness. This case exhibits very late immune response or aggressive growth of antigen leading to high antigen concentration consuming most of the resources of the organism thus causing significant pathological changes in an organism and consequent failure to sustain or intensify the immune response. In either case, the multiplication rate of the specialized immune cells becomes insufficient for the suppression of antigen.

The simulation study conducted on the preliminary model of the immune response [1] facilitates convincing visualization of the complex phenomena of the immune response (see below). It could be seen that in the first case the lethality is caused by a high volume of the antigen introduced in the organism, and in the second case by low resource level of the organism in combination with low initial immunity. It should be noted that the moment of death is recognized as the instance when the resource level of the organism becomes equal to zero. Note the gradual decline of the antigen after the death of the host.

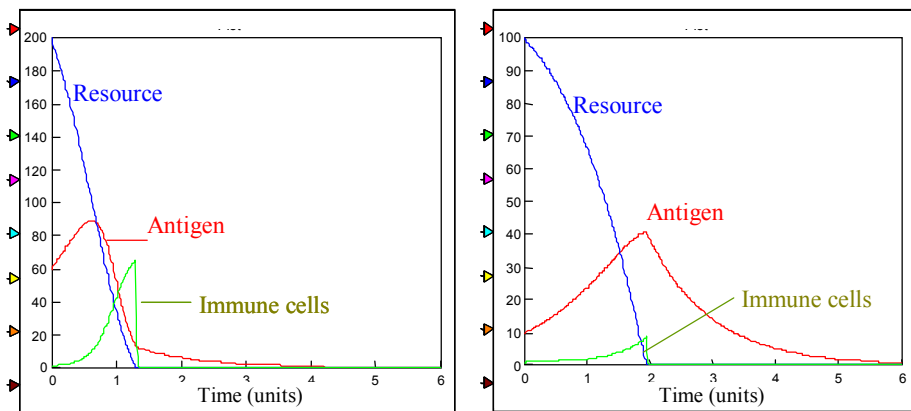


Figure 3 – Immune responses resulting in lethality.

The chronic illness outcome of an immune response could be also easily modeled and visualized using the existing model as a situation when the rate of multiplication of the antigen and immune cells and the restoration of the resources of the organism reach equilibrium. This situation is well known in immunology and medicine.

It should be emphasized that three major components of the immune response, the antigen growth, the growth of the immune cells, and the dynamics of the organism resources for an obvious closed-loop chain of transformations that constitutes a highly nonlinear, time-dependent, multivariable negative feedback system that *does confirm* to advanced analysis and synthesis theory developed in modern controls.

The refinement of the presented model and detailed analysis of specific situations known in modern immunology are being conducted. However, the described model cannot be fully exploited in immunology. It is difficult to assign values to the various parameters and to adjust the responses. Many of the parameters are locked in the complexity of the immune system and parameters are simple fundamentals of nature. In a system of computer networks, at any scale, each parameter can be estimated. Simulations would provide network planners with an accurate estimate on how to harden networks and services. Furthermore, unlike a biological system, the response can be tuned to provide the desired outcome.

4 Control of the immune response

The resultant mathematical model of the immune response facilitates the understanding and numerical simulation of major interactions between the components of the immune system and the intruding antigen. Subjected to analytical techniques offered by modern control theory and numerical analysis, this model enables us to establish the conditions for three possible outcomes of such interaction, full recovery, chronic infection, and lethality, and to formulate a control law “enforcing” the full recovery outcome. The control law developed for the closed-loop system represented by the model is visualized as a mechanism capable of modifying the properties of the underlying biological phenomena in order to assure the full recovery outcome. It could be seen that the full recovery outcome can be assured by

- high initial immunity (i.e. the enhanced ability to detect the intruding antigen),
- high rate of multiplication of the specialized immune cells, and
- maintaining the resource level of the organism necessary to sustain the high rate of multiplication of the immune cells. Since these characteristics are defined by the appropriate numerical values of the parameters of the mathematical model, the control law could be expressed by the definition of the respective model parameters as functions of the current level of the antigen concentration. Unfortunately, this control law implies the modification of major biological mechanisms and as such cannot be recommended to a physician and can be applied only to the simulation model of the immune response.

However, the main value of this model justifying the effort of its development is that it provides a description of a highly successful active response of a complex system utilizing a very large number of semi-autonomous agents. Consequently, describing defenses of a computer network against information attack in terms of this model provides the means for the analysis of this complex process and allows for the adoption of the established control law leading to the only desirable outcome, the full recovery. Unlike the immune system, characteristics of a man-made computer network are very likely to be adequately represented by the model parameters.

5 Anatomy of an information attack on a computer network. A recent example

The following chronology of events, describing one of the recent internationally known information attacks, is presented to emphasize its similarity to the chain of events observed in the immune system defending the biological organism against antigen.

In August and September of 2003, the W32/Blaster worm and its variants targeted computer systems running various Microsoft Windows operating systems. Before the onslaught, there came plenty of warnings with recommended actions for system administrators and computer owners to prevent potential infection from a yet unknown (and non-existent) assailant. On July 16, 2003, Microsoft released a security bulletin (MS03-026) with a *critical* severity rating and a software patch to update affected systems. According to Microsoft, the security hole would allow a remote potential assailant to “run code of attacker’s choice.”

On July 17, 2003, many news outlets covered the newly discovered vulnerability as the first discovered flaw in the Windows Server 2003 platform and indicated that previous versions of the operating system were also affected. Microsoft indicated that this security hole would allow for hackers to seize control of the target computer, steal data, damage files, or potentially eavesdrop on communications [Fox News July 17, <http://www.foxnews.com/story/0,2933,92212,00.html> - “Microsoft Admits Flaw in Windows Software”]. People running the installed operating systems were urged to install the software patch to correct the problem.

August 11, 2003 was when the first attack from the W32/Blaster worm was detected. CERT, a well established center for reporting software vulnerabilities and virus/worm related incidents, issued advisory CA-2003-20 detailing the discovery, but offered little advice on how to eliminate infection or prevent infection on non-patched systems. Headlines followed

with Fox News reporting “Internet Attack Disables Thousands of Computers” on August 13 and the Department of Homeland Security issuing an Internet worm advisory on August 14 [http://www.dhs.gov/interweb/assetlibrary/Blaster_Worm_Advisory.PDF]. Within three days of the initial detection, the major anti-virus companies issued updates to their virus detection software to detect Blaster, CERT, Microsoft, and the Department of Homeland Security had listed methods and techniques to prevent infection or to repair a compromised system, and over 188,000 computer systems worldwide were infected [Network World Fusion <http://www.nwfusion.com/news/2003/0812blastinfect.html> - “Update: Blaster worm infections spreading rapidly”]. Even though the vulnerability was known for approximately a month before the actual infection, severe damage occurred. High profile users, such as the Maryland Department of Motor Vehicles and the U.S. State Department, were among its victims. Soon after the release of Blaster, the mutants, using similar means for infection, began to propagate.

The Blaster worm infects a host by exploiting a buffer overrun in Microsoft’s DCOM RPC service. A carefully crafted packet is sent to a host using TCP port 139. This packet results in a buffer overflow that causes code transferred through this connection to begin executing. Upon execution, a copy of Blaster is obtained from the infecting host and is installed on the newly infected machine. At this point, the infected host executes Blaster and begins to infect other hosts using the same technique. Part of the payload of the worm is a TCP SYN flood against the very servers at Microsoft that distributed the needed software patch to permanently immunize a system from Blaster.

Before the assault, except for the systems that had received the corrective patch or were behind properly configured firewalls or were naturally immune to Blaster, there was no mechanism to detect an attempt to infect any given target machine. Detection of the attack only occurred after infection by the computer operator and only after the new host machine spread the infection to a number of other hosts. In other words, the initial concentration of immune detectors was quite small. As more machines became infected, the concentration of antigen increased. This continued until computer operators noticed unusual activity, reported it and the response began. This is referred to as the threshold of detection.

As Blaster spread, available resources were consumed. Ultimately, these resources were uninfected vulnerable hosts, bandwidth to the Microsoft update servers, bandwidth of compromised machines and the number of operators available to respond to the infected machines. Furthermore, already infected machines became victims of Denial of Service (DoS) attacks caused by future attempts to infect. These machines typically stopped responding or forcibly reboot.

The response began with organizations releasing information about the new worm and methods to prevent infection. Antivirus manufacturers released updates to their definition files that would detect and stop Blaster from infecting a vulnerable host or prevent a host from infecting future machines, however, the “protected” machine would still hang or reboot after an unsuccessful attempt at infection, preventing further legitimate use of the system. The installation of the updated definition files ultimately rested in a valid subscription to the definition files or a specific action on behalf of a computer operator. Furthermore, Microsoft released information on how to prevent infection of an unpatched host. The consequence of the prevention measures was that some heavily used remote services would become crippled.

With each newly developed or implemented defense technique, the effective concentration of fighter cells increased. The cost of the response included both manpower and bandwidth. Each infected machine required an operator to physically recover it. The recovery process involved obtaining the patch from Microsoft, requiring bandwidth, and a period of time where the machine, while obtaining and installing the patch didn’t malfunction as a result of another attack.

Within a few weeks of the first infection, hundreds of thousands of machines were affected. The response took several days, during which mutants of Blaster began circulating. Many machines became permanently immune during this period, with the proper software patches installed. However, many machines existed in “safe” environments, out of the reach of Blaster, until some event in the

future re-exposes them to risk. As an example, one of the responses to the worm was to firewall off the network services that Blaster exploited. It can be realized that these are only temporary, and in the future, the firewall protection can be eliminated. This slow denigration of defenses is one of the components of the natural aging of the computer network.

The primary reason that Blaster was able to perform as much damage as it did is a direct result of a long delay period between the initial infection and the response. A large number of machines became infected consuming a large quantity of resources that could have otherwise gone into fighting off the worm. This results from a primarily *human* detection mechanism and a *human* protection mechanism. After all, if all of the systems were inoculated (with the software patch), there would not have been an infection. Or, if an alternative mechanism could have challenged the initial infection similar to ones proposed in [5] and [8], this delay period could have been seriously reduced.

6 The computer network considerations

The parameters obtained for the model of the immune system can be extended to computer systems and networks of computer systems to model the effects of various attack mechanisms. The prime difference, however, between a computer network and a living organism is that each of these parameters needs to be established for a particular attack mechanism and a specific service or group of services. The mechanisms of immune responses described by the immune model are relatively uniform, with many of the same cellular groups performing common functions. Furthermore, resource availability is relatively uniform for any given attacker and any given response. For consistency, extension of a biological immune response model to a specific information attack on a computer network will require a specific set of parameters.

The level of available resources within a computer network can be comprised of many different elements and referred to as the *throughput* of the computer network. This is an aggregate parameter, composed of at least one and potentially many parameters. Within a computer network, there are a number of components that can be consumed by both that attacker and the response mechanism. Examples of these components are processing capability of the various networking components, bandwidth, queue depths, machine and process throughput and latency, etc.

A simple example is a TCP/IP SYN flood attack, which is a form of denial of service (DoS) attack. The point of the assault is to render the server attached to the network *offline*. This is accomplished by filling the half-open TCP connection queue with phony connection requests. The important components of the network resources are the TCP connection queue depth and CPU time required to analyze each connection to verify authenticity of the host attempting the connection. A more complicated example is an Internet worm, the likes of Code Red or W32.Blaster, which, through their replication and subsequent infection disable host after host. In this case, the required resources is composed of valid “uninfected” hosts, human resources available to apply the appropriate patches, network throughput for the spread of infection, and time to apply the appropriate fix before the host computer is disabled.

The relative concentration of *detector cells and responders* for the attack mechanism can be represented by the ratio of computer operators to the number of machines on a network with properly configured and updated anti-virus software to a complex mechanism analyzing every packet exchanged on a network. As in the biological immune system, many of the detection mechanisms in the computing world are forms of signature detection. As in a biological system, the detectors are free to roam and detection occurs when the detector happens upon an antigen, to reduce overhead, network packets may be sampled or only checked against a specific signature and allowed to pass. In this case, the detector concentration can simply be the number of signatures within a large set applied to each packet at a specific checkpoint.

The concentration of the antigen represents, in computer terms, the frequency or depth of the attack. As in the biological immune system, this is an attack specific parameter. This can be the number

of machines infected with a computer virus, the number of half open TCP/IP SYN requests per second, level of bandwidth consumption by malicious activity, number of e-mail messages processed per second, etc.

In an immune system, there is a parameter that describes the threshold of detection, which represents the concentration of antigen at which detection occurs. This parameter is inversely proportional detector concentration. For example, if a virus or worm attacks and is previously unknown, at a minimum, it is simply related to the number of trained computer operators which directly observe the attack.

The coefficients of an organism's *recovery rate* represent the restoration of resources as a result of immune activity and due to normal recovery. There is no simple analog to this parameter in a computing network. This can be viewed as a complex interaction of computer administrators ensuring clean and well-protected systems along with properly sized and maintained network connections.

The *ageing of the immune memory* ultimately relates to the *human factor* of the network. At any given point in time, a computer system may be considered well protected, however, in the passage of time, various exploits in installed software are uncovered, holes in networking protocols are exposed and new viruses are developed. Unless the network is continuously undergoing updates, these weaknesses accumulate until some later event triggers an update. This phenomenon was demonstrated by the heavily publicized attack from the W32.Blaster worm. A patch, which repaired a previously discovered software deficiency, was available a month before the worm infected its first host.

During an attack and the subsequent immune response, the concentration of antigen undergoes cascade increase. Only after detection does the immune response begin. At this point systems administrators begin the first portion of the response, isolating infected machines, reinforcing existing defense mechanisms, isolation of *clean* networks to preserve integrity, installation of detection and repair software and infected systems recovery. Once mobilized, the response also undergoes cascade increase and continues until all affected systems are recovered and computer network systems performance is recovered.

It can be argued that in a computer network, ultimately, resources to repair and restore a network system are unlimited. Money and manpower are the prime resources and in the end will always succeed. Therefore, *death of the network* does not happen in the same sense as a biological organism. In a biological organism, when death occurs, it is final. In computing systems, true death is only the *permanent loss of information*. However, serious disruptions in computing services in high demand time-sensitive applications can and do cause loss of equipment, valuable resources and even human lives.

Knowing how to relate the parameters of a computer network to the biological immune system allow for modeling of various attack, detection of bottlenecks and assessment of vulnerabilities of the computer network. Furthermore, deficiencies in current network defenses can also be identified.

The control law established for the model of the immune response could be reformulated in terms of parameters of the computer network and implemented in automatic regime and/or through the actions of the network manager, providing a numerically justified basis for the reconfiguration of the network subjected to the information attack.

References

- [1] V. Skormin, "Biological Approach to System Information Security (BASIS): A New Paradigm in Automatic Information Assurance". Final Report to AFRL at Rome NY on Contract #30602-01-0509, Binghamton, NY, 2002.
- [2] Skormin, V.A., Delgado-Frias, J.G., McGee, D.L., Giordano, J.V., Popyack, L.J., Gorodetski, V.I. and Tarakanov, A.O., BASIS: a biological approach to system information security, *Information Assurance in Computer Networks* (Gorodetsky V.I., Skormin V.A. and Popyack L.J. eds. LNCS 2052, Springer-Verlag, Berlin, 2001, pp. 127-142).

- [3] V. Skormin, D. Summerville, J. Moronski, D. McGee, "[Biological Approach to System Information Security \(BASIS\): A Multi-agent Approach to Information Security](#)", Lecture Notes in Computer Science, Volume 2691, Springer-Verlag Heidelberg, 2003.
- [4] A. Tarakanov, V. Skormin, S. Sokolova, "Immunocomputing. Principles and Applications", 210 pp, Springer-NY, 2003.
- [5] V. Skormin, D. Summerville, J. Moronski, "Detecting Malicious Codes by the presence of their *Gene of Self-Replication*", "Computer Network Security", Lecture Notes in Computer Science, Volume 2776, Springer, 2003.
- [6] V. I. Gorodetski, I. V. Kotenko, L. J. Popyack, and V. A. Skormin, "Agent-Based Model of Information Security System: Architecture and Framework for Behavior Coordination", *Proceedings of the First International Workshop of Central and Eastern Europe on Multi-agent Systems (CEEMAS'99)*, June 1999, pp. 323-331.
- [7] V. Gorodetski, V. Skormin, L. Popyack (Eds.), "Information Assurance in Computer Networks. Methods, Models and Architectures for Network Security", Lecture Notes in Computer Science, Springer, 2001.
- [8] V. Skormin, D. Summerville, J. Moronski, and J. Sidoran, "Application of Genetic Optimization and Statistical Analysis for Detecting Attacks on a Computer Network", *Proceedings of the Real-time Intrusion Detection NATO Symposium*, May 27-29, Lisbon, Portugal, 2002.
- [9] V. Gorodetski, L. Popyack, V. Skormin (Eds.), "Computer Network Security", Lecture Notes in Computer Science, Springer, 2003.
- [10] T. Gard and J. Hoffacker, "Asymptotic Behavior of Natural Growth on Time Scales", *Dynamic Systems and Applications*, Vol. 11, 2002.
- [11] Anderson RW, Neumann AU, Perelson AS (Theoretical Biology and Biophysics, Los Alamos National Laboratory). A Cayley tree immune network model with antibody dynamics // *Bull Math Biol* 1993 Nov.
- [12] Singer DF, Linderman JJ (Department of Chemical Engineering, University of Michigan, USA). The relationship between antigen concentration, antigen internalization, and antigenic complexes: modeling insights into antigen processing and presentation // *J Cell Biol* 1990 Jul.
- [13] Seiden PE, Celada F. (IBM T. J. Watson Research Center) A model for simulating cognate recognition and response in the immune system // *J Theor Biol* 1992 Oct.
- [14] Merrill S. Mathematical models of Humoral immune response // *Techn. Rep. of the Univ. Of Iowa*, 1976. – 40 p.
- [15] Batt BC, Kompala DS (Department of Chemical Engineering, University of Colorado). Verification of immune response optimality through cybernetic modeling // - *J Theor Biol* 1990 Feb.
- [16] Singer DF, Linderman JJ (Department of Chemical Engineering, University of Michigan, USA). The relationship between antigen concentration, antigen internalization, and antigenic complexes: modeling insights into antigen processing and presentation // *J Cell Biol* 1990 Jul.
- [17] Weinand RG, Conrad M (Department of Computer Science, Wayne State University, Detroit, Michigan). Maturation of the immune response: a computational model // *J Theor Biol* 1988 Aug 22.
- [18] Morel PA (University of Pittsburgh). Mathematical modeling of immunological reactions // *Front Biosci* 1998 Mar.
- [19] Mohler R., Barton C. Compartmental control model of the immune process: *Proceedings of the 8th IFIP Conference on Optimization Techniques*. – Heidelberg: Springer-Verlag, 1978.
- [20] Mohler R. Bilinear control structures in immunology // *Proc. Of IFIP Working Conference on Modelling and Optimization of Complex systems*. – Berlin a.o.: Springer-Verlag, 1979.
- [21] Mohler R., Barton C., Hsu C. System theoretic control in immunology. – Oregon State Univ., 1975.
- [22] S. Forrest, F. Hofmeyr, A. Somayaji, T.A. Longstaff, "A Sense of Self for Unix Processes", *Proceedings of the 1996 IEEE Symposium on Security and Privacy*, IEEE Computer Society Press, Los Alamitos, CA, 1996.
- [23] D. Dasgupta (Ed.) *Artificial Immune Systems and Their Applications*, Springer, 1998.
- [24] D. Dasgupta and F. Gonzalez. An Immunity-Based Technique to Characterize Intrusions in Computer Networks. To appear in the journal *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 3, June 2002.

МЕТОДЫ РЕШЕНИЯ ЗАДАЧИ ТИПАЖА СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ

Б.П. Топоров, В.М. Горлов

ФГУП ГосНИИ авиационных систем
123319, Москва, ул. Викторенко, 7, Россия
vmgorlov@mail.ru
тел: +7 (8095) 157-92-78

Ключевые слова: сложная техническая система, облик, типаж, оптимизация, структура.

Abstract

For the solution of a common problem "synthesis-type" we shall consider actually the task of the type of complex technical systems (CTCs), when the set of CTC alternative types with known characteristics of the external layout is preset, and it is required to generate from them rationally a park which is capable to execute the preset combination of target tasks. Such task arises always during the solution of the synthesis CTC common problem on one of the stages (steps) of the in-house iterative cycle, and frequently in independent practical cases, when it is required to give recommendations on the expediency of developing a new CTC type, proposed for creation, or on continuation of manufacture and operation of CTC already in operation. The methods of solving the task of type are proposed.

1 Постановка задачи

Для решения общей проблемы "синтез-типаж" [1] рассмотрим собственно задачу типажу сложных технических систем (СТС), когда задан набор альтернативных типов СТС с известными обликowymi характеристиками и требуется из них сформировать рациональным образом парк, способный выполнить заданную совокупность целевых задач S . Очевидно, что множество S должно поступать на вход модели с более высоких иерархических уровней, например, для СТС, как результат моделирования (в том числе и эвристического) операции использования этих систем, и может корректироваться заказчиком по результатам решения задачи "синтез-типаж".

Выбор критерия оптимальности представляет собой самостоятельную проблему, как известно, неподдающуюся формализации. Сложившиеся в теории системного анализа и исследования операций общие подходы к этой проблеме диктуют среди немногих обязательных требований к критерию требование достаточно полного формального отражения существенных связей оптимизируемой системы с системами более высокого иерархического уровня. На этом уровне определяются целевые задачи множества S . В свою очередь, оптимизируемая система отображается в модели верхнего иерархического уровня рядом параметров, которые и должны, по существу, лечь в основу при формировании критерия оптимальности.

Для конкретных исследований по оптимальному синтезу СТС в качестве критерия оптимальности можно рекомендовать использование стоимости C_{Σ} потребного парка СТС, добиваясь ее минимума при условии выполнения всей заданной совокупности $S(t)$ целевых задач парка

$$(1) \quad C_{\Sigma} = \sum_k^K C_{\Sigma_k} \rightarrow \min \text{ при } S(t) \geq S^*(t),$$

где $S^*(t)$ - заданная совокупность целевых задач, решаемых парком СТС;

$$C_{\Sigma_k} = \sum_{i=1}^{N_k(t)} C_{k_i}(t, N_k) + C_{k_0}(t) + C_{k_s}(t, N_k, t_s) + C_{k_{кан}}(t, N_k);$$

$C_{k_i}(t, N_k)$ - стоимость i -ого образца СТС k -го типа при производстве серии размером N_k ; N_k - суммарное число СТС k -го типа (всего типов K), входящих в парк;

$C_{k_0}(t)$ - затраты на разработку (НИОКР) СТС k -го типа;

$C_{k_s}(t, N_k, t_s)$ - затраты на эксплуатацию парка N_k СТС k -го типа в течение времени t_s ;

$C_{k_{кан}}(t, N_k)$ - затраты на капитальное строительство.

Если определено множество целевых задач, решаемых СТС k -го типа $s_k(t) = \{n_{j_k}(t, l, \tau), j \in J, 0 \leq l \leq L, 0 \leq \tau \leq T\}$ такое, что $\bigcup_{k=1}^K s_k(t) = s(t)$, то потребное для их выполнения суммарное число N_k СТС можно, очевидно однозначно, определить, зная технические возможности каждой СТС, определяемые ее обликowymi характеристиками.

Располагая информацией о целевых задачах, возможных альтернативах СТС, можно приступить к решению задачи определения рационального типажа СТС. При этом все возникающие ситуации могут быть отнесены к одной из следующих типовых задач.

I. Статическая задача типажа (все рассматриваемые величины не зависят от времени t): Задан суммарный объем целевых задач S , который должен быть решен формируемым парком СТС в ходе ряда типовых операций. Задан набор альтернативных вариантов СТС X , причем технические, эффективностные и экономические характеристики этих СТС известны.

Необходимо определить рациональный состав парка СТС, обеспечивающий выполнение заданного объема целевых задач при минимальных суммарных затратах на разработку, серийное производство и эксплуатацию в течение принятого интервала планирования Δt (или в течение полного жизненного цикла СТС) всех входящих в парк СТС.

II. Статическая задача типажа с памятью: В отличие от задачи I, где все СТС считаются вновь разрабатываемыми, здесь предполагается, что к моменту формирования рационального парка в эксплуатации имеется набор Y СТС старых типов, производство которых уже прекращено или еще продолжается. В соответствии с этим при вычислении критериальной функции (суммарных затрат) для этих СТС учитываются только затраты на эксплуатацию и серийное производство (если оно продолжается).

III. Динамическая задача типажа: Представляет собой динамический (во времени планируемого периода) аналог задачи II. Постановка задачи III помимо чисто формального введения явной зависимости от времени всех величин, входящих в описание задачи, означает переход от задачи многопараметрической оптимизации к неклассической вариационной задаче и требует указания начала активного использования СТС (т.е. t_0 - момент реформирования парка СТС) [1]. Рассмотрим последовательно методы решения сформулированных задач.

2 Решение задач

2.1 Метод решения статической задачи типажа (задача I)

Как известно, удачный выбор единиц измерения для параметров модели может существенно упростить решение задачи. В частности, оказывается удобным и в большинстве случаев возможным измерять парк СТС и возложенный на него объем целевых задач в одних и тех же единицах. Это позволяет привести задачу I к стандартному виду задачи математического программирования с линейными ограничениями. В качестве таких единиц измерения можно выбрать, например, эквивалентные эталонные СТС [2]. Для каждой целевой задачи может быть определено потребное количество СТС данного типа. Если рассмотреть стандартную типовую

задачу, выполняемую СТС, то можно пересчитать полученное потребное количество в количество СТС, выполняющих минимально целесообразный объем с учетом соответствующих потерь, если решается конфликтная задача. Подобный пересчет позволяет записать следующее уравнение баланса: $\sum_{i=1}^n s_i \leq \sum_{i=1}^n \sum_{j=1}^K x_{ij}$, где s_i - количество подлежащих выполнению целевых задач i -го типа (всего n типов целевых задач); x_{ij} - количество СТС j -го типа (всего K типов СТС), выделяемых для решения целевых задач i -го типа.

Общее уравнение баланса является следствием частных балансных уравнений:

$$(2) \quad s_i \leq \sum_{j=1}^K x_{ij}, \quad i \in \overline{1, n}.$$

В качестве первого приближения к модели типажа СТС (задача I) может быть сформулирована следующая транспортная задача линейного программирования:

найти вектор x , доставляющий минимум линейной форме

$$(3) \quad \sum_{j=1}^K C_j \sum_{i=1}^n x_{ij}$$

при ограничениях (2) и

$$(4) \quad x_{ij} \geq 0, \quad i \in \overline{1, n}, \quad j \in \overline{1, K},$$

где C_j - суммарные затраты на разработку, производство и эксплуатацию одного СТС j -го типа. Эта задача является вырожденной, поскольку в ней отсутствуют ограничения вида $\sum_{i=1}^n x_{ij} \leq x_j \geq 0, \quad j \in \overline{1, K}$, то есть ограничения на "емкость складов", и имеет очевидное решение:

$x_{ij} = \begin{cases} 0, & j \neq r \\ s_i, & j = r \end{cases} \quad i \in \overline{1, n}, \quad j \in \overline{1, K}$, где единственный индекс r находится из соотношения

$C_r = \min_{j \in \overline{1, K}} C_j$. Полученное решение не является тривиальной рекомендацией комплектовать

парк из единственного наиболее дешевого типа СТС, поскольку в балансных уравнениях (2) некоторые x_{ij} являются заведомо нулевыми, это значит, что СТС соответствующего j -го типа не в состоянии выполнить задачу i -го типа, например, в силу недостаточной мощности системы. Полученная структура оптимального решения позволяет также оценить возможность изменения рационального типажа СТС при вариации целевых задач. Введение $(n+1)$ -ой задачи, очевидно, не может повлечь за собой исключение одного из типов СТС из рационального парка. При этом включение m -го типа СТС в рациональный парк возможно, если выполняется соотношение $C_m x_{n+1, m} = \min_k C_k x_{n+1, k} < C_l x_{n+1, l}$, где l - индекс типов СТС, уже включенных в

парк. Аналогично можно записать условие включения $(K+1)$ -го типа СТС в ранее сформированный рациональный парк; при этом должно выполняться хотя бы одно из следующих соотношений: $C_K x_{K+1, i} = \min_{\eta \in \overline{1, K}} C_\eta x_{\eta+1, i}, \quad i \in \overline{1, n}$. Модель (3)-(4) представляет тем не менее модель

первого (если не нулевого) приближения, поскольку в ней не учитывается наличие очевидной зависимости $C_j(x_j)$, обусловленной проявлением эффекта серийности (зависимости затрат на

СТС от размеров его серийного производства). Соответствующая экспликация модели (3)-(4) приводит к следующей задаче нелинейного математического программирования:

найти вектор x , доставляющий минимум функции

$$(5) \quad C_{\Sigma} = \sum_{j=1}^K C_{\Sigma_j}(x_j) = \sum_{j=1}^K C_j(x_j) \sum_{j=1}^K x_{ij}$$

при ограничениях (2), (4).

Сформулируем в виде лемм несколько положений, опуская доказательства.

Лемма 1. Целевая функция (5) вогнута по переменным x_j .

Теперь нетрудно видеть, что задача (5) относится к классу задач сепарабельного программирования, и при ее решении может быть использована одна из вычислительных процедур метода динамического программирования. При этом сформулированное ниже предложение позволяет существенно сократить объем необходимых вычислений.

Лемма 2. Оптимальное решение задачи (5) имеет вид

$$(6) \quad x_{ij} = \begin{cases} 0, & j \neq r \\ s_i, & j = r \end{cases} \quad i \in \overline{1, n}, \quad j \in \overline{1, K}.$$

Другими словами, использование СТС различного типа для решения одной и той же целевой задачи нецелесообразно. Лемма 2 оказывается весьма полезной, поскольку она является основанием для отказа от рассмотрения смешанных стратегий при использовании СТС для решения отдельных задач. Введем следующее специальное упорядочение целевых задач СТС. Будем считать, что элементы множества S ранжированы по степени возможности их выполнения различными СТС от простых к сложным. В частности, задачу, получившую индекс $i=1$, в состоянии выполнить все альтернативные варианты СТС, а задачу, получившую индекс $i=n$, только несколько наиболее сложных и дорогих из рассматриваемых типов СТС. Введение упорядочения целевых задач порождает и соответствующее разбиение типов СТС на n классов. При этом СТС, принадлежащие i -му классу (они могут быть и различных типов), в состоянии решать любую j -ю задачу при $j \leq i$ и не в состоянии сделать это при $j > i$. Как следует из доказанной леммы 2 в рациональный парк войдут только по одному представителю от каждого класса СТС, поэтому в дальнейшем вместо обозначения x_{ij} будет использоваться обозначение x_i . Причем, конечно, не исключается, что для решения различных задач в рациональном парке будут использоваться СТС одного и того же типа. В этих предположениях из частных балансных уравнений может быть получена система $\sum_{i=j}^n x_i \geq \sum_{i=j}^n s_i$, $j \in \overline{1, n}$, или в развернутой форме

$$\begin{aligned} x_1 + x_2 + x_3 + \dots + x_{n-1} + x_n &\geq s_1 + s_2 + s_3 + \dots + s_{n-1} + s_n, \\ x_2 + x_3 + \dots + x_{n-1} + x_n &\geq s_2 + s_3 + \dots + s_{n-1} + s_n, \\ x_3 + \dots + x_{n-1} + x_n &\geq s_3 + \dots + s_{n-1} + s_n, \\ &\dots \\ x_{n-1} + x_n &\geq s_{n-1} + s_n, \\ x_n &\geq s_n. \end{aligned}$$

В соответствии с принятой в динамическом программировании процедурой введем наряду с рассматриваемой задачей I последовательность задач $I^{(k)}$, $k \in \overline{1, n}$. Каждая из задач $I^{(k)}$ формулируется следующим образом: найти рациональный парк СТС $\tilde{x}^{(k)} = (\tilde{x}_1^{(k)}, \dots, \tilde{x}_k^{(k)})$, выполняющий заданный объем целевых задач $s^{(k)} = (s_1 + s_2 + \dots + s_k)$ при минимальном значении

целевой функции $C_{\Sigma} = \sum_{j=1}^K C_{\Sigma_j}(\tilde{x}_j^{(k)})$. Понятно, что задачи $I^{(n)}$ и I совпадают. Сформулируем теперь основной результат, позволяющий построить простой шаговый алгоритм нахождения решения задачи I .

Лемма 3. Оптимальный вектор \tilde{x} в задаче I является решением функционального уравнения $C_{\Sigma}(\tilde{x}) = \min C_{\Sigma}(\tilde{x}^{(k-1)}, 0, \dots, 0, s_k + \dots + s_n)$.

Другими словами, рациональный парка \tilde{x} представляет собой один из вариантов $(\tilde{x}^{(k-1)}, 0, \dots, 0, s_k + \dots + s_n)$, $k \in \overline{1, n}$. Лемма 3 сводит задачу I к решению последовательности задач $I^{(1)}, I^{(2)}, \dots, I^{(n-1)}$. В свою очередь, задача $I^{(n-1)}$ сводится к последовательности задач $I^{(1)}, I^{(2)}, \dots, I^{(n-2)}$. Теперь может быть формально описан алгоритм (типа вычислительных процедур динамического программирования), который дает решение задачи I за n шагов.

Алгоритм последовательного подъема:

Первый шаг. Решается задача $I^{(1)}$: $\tilde{x}^{(1)} = s_1$, $c_1 = c_{\Sigma_1}(s_1)$.

Второй шаг. Решается задача $I^{(2)}$: $c_2 = \min[c_1 + c_{\Sigma_2}(s_2), c_{\Sigma_2}(s_1 + s_2)]$. Имеем $\tilde{x}^{(2)} = (\tilde{x}^{(1)}, s_2)$, если $c_2 = c_1 + c_{\Sigma_2}(s_2)$ или $\tilde{x}^{(2)} = (0, s_1 + s_2)$, $c_2 = c_{\Sigma_2}(s_1 + s_2)$.

i -й шаг. Решается задача $I^{(i)}$: $c_i = \min_{k \in \overline{0, i-1}} [c_k + c_{\Sigma_i}(s_{k+1} + \dots + s_i)]$. Здесь принято, что

$C_0 = 0$. Пусть искомым минимум достигается при $k = k_i$. Тогда $\tilde{x}^{(i)} = (\tilde{x}^{(k_i)}, 0, \dots, 0, s_{k_i+1} + \dots + s_i)$.

n -й шаг. Решается задача $I^{(n)} = I$: $c_n = \min_{k \in \overline{0, n-1}} [c_k + c_{\Sigma_n}(s_{k+1} + \dots + s_n)]$. Здесь принято, что

$C_0 = 0$. Пусть искомым минимум достигается при $k = k_n$. Тогда $\tilde{x} = (\tilde{x}^{(k_n)}, 0, \dots, 0, s_{k_n+1} + \dots + s_n)$. Этот алгоритм может использоваться и при решении более сложных задач II и III .

2.2 Метод решения задачи II

В ранее сделанных предположениях условия выполнения парком СТС заданного объема целевых задач записываются в виде следующей системы балансных уравнений

$$y_1 + x_1 + y_2 + x_2 + \dots + y_n + x_n \geq s_1 + s_2 + \dots + s_n,$$

$$y_2 + x_2 + \dots + y_n + x_n \geq s_2 + \dots + s_n,$$

$$y_n + x_n \geq s_n.$$

Кроме того, в соответствии с физическим смыслом задачи должны выполняться неравенства

$$(7) \quad x_i \geq 0, \quad 0 \leq y_i \leq \mathcal{G}_i.$$

Задача II представляет собой следующую задачу нелинейного математического программирования: найти векторы x, y , доставляющие минимум функции

$$C_{\Sigma} = \sum_{j=1}^n C_{\Sigma_j}(x_j) + \sum_{j=1}^n C_{\Sigma_j}(y_j) \text{ при указанных выше ограничениях. В математическом плане задача } II \text{ существенно сложнее задачи } I \text{ в том отношении, что структура целевой функции уже}$$

$s^{<k>} = (s_k + s_{k+1} + \dots + s_n)$ при минимальном значении целевой функции $C_{\Sigma}^{<k>} = \sum_{j=k}^n C_{\Sigma_j}(\tilde{z}_j^{<k>} y_j)$. Понятно, что задачи $\Pi^{(n)}$, $\Pi^{<1>}$ и Π совпадают. Формальное обоснование алгоритма решения задачи Π дает следующая лемма (доказательство опускаем).

Лемма 5. Рациональный парк СТС \tilde{z} в задаче Π совпадает с одним из n следующих вариантов: $(\tilde{z}^{<n-1>}, \tilde{z}^{<n>}), \dots, (\tilde{z}^{<k-1>}, \tilde{z}^{<k>}), \dots, (\tilde{z}^{<1>}, \tilde{z}^{<2>})$, где $\tilde{z}^{<k>}$ находится как решение функционального уравнения $C_{\Sigma}^{<k>}(\tilde{z}^{<k>}) = \min_{t \in k, n} C_{\Sigma}^{<k>}(\tilde{z}_t^{<k>})$.

Следует отметить, что парк СТС $\tilde{z}^{<k>}$ обращает все балансные уравнения, кроме k -го, в строгие неравенства. Используя этот факт, можно доказать лемму 6.

Лемма 6. Среди компонент $\tilde{x}^{<k>}$ вектора $\tilde{z}^{<k>}$ имеется не более одной положительной.

На основании леммы 6 может быть предложена следующая процедура отыскания вектора $\tilde{z}^{<k>}$. Удалим из вектора $\tilde{z}^{<k>}$ все компоненты x_i за исключением одной x_t . Решение альтернативной задачи $\Pi_t^{<k>}$ может быть легко получено с помощью выше приведенного алгоритма. Заметим, что если для парка $\tilde{z}^{<k>}$ выполняются балансные уравнения (7), то он совпадает с одним из вариантов парка $\tilde{z}^{<k>}$, $t \in k, n$, а именно с тем из них, для которого справедливо функциональное уравнение $C_{\Sigma}^{<k>}(\tilde{z}^{<k>}) = \min_{t \in k, n} C_{\Sigma}^{<k>}(\tilde{z}_t^{<k>})$.

2.3 Задача III - динамическая задача типажа СТС

В качестве метода решения этой задачи предлагается J раз (J - число интервалов планирования в рассматриваемом периоде времени) применять изложенные выше методы к последовательности статических задач Π_i , в которых в качестве исходных данных, описывающих достигнутое состояние парка, используются результаты решения предыдущей задачи Π_{i-1} . Такой подход позволяет получить субоптимальное решение. В качестве границ подинтервалов целесообразно рассматривать возможные даты поступления в эксплуатацию альтернативных типов СТС.

Заметим, что модели рационального типажа могут рассматриваться как математический аппарат, дающий полезную информацию для принятия соответствующих решений, а их применение может быть продуктивным только в случае инженерного анализа, как используемых исходных данных, так и получаемых результатов.

Список литературы

- [1] Топоров Б.П., Горлов В.М. "Проблемы формирования рационального облика и типажа сложных целевых систем, функционирующих в условиях конфликта и неопределенностей". Труды IV Международной конференции "Проблемы управления и моделирования в сложных системах", 2002 г.
- [2] Топоров Б.П., Горлов В.М. "Итеративная обратимая инвариантная эталонизация в многомерных динамических системах". Труды II Международной конференции "Проблемы управления и моделирования в сложных системах", 2000 г.
- [3] Жеребин А.М., Топоров Б.П., Горлов В.М. "Методы моделирования в задачах анализа и управления при проектировании авиационных комплексов". Труды III Международной конференции "Проблемы управления и моделирования в сложных системах", 2001 г.

МЕТОДЫ СТРУКТУРНОЙ ТЕОРИИ В ЗАДАЧАХ СИНТЕЗА СИСТЕМ УПРАВЛЕНИЯ С РАСПРЕДЕЛЕННЫМИ ПАРАМЕТРАМИ

Э.Я. Рапопорт

Институт проблем управления сложными системами РАН
443020, г. Самара, ул. Садовая, 61, Россия
тел. (8462) 32-26-77, факс (8462) 33-27-70

Ключевые слова: структурная теория, синтез, алгоритмы управления, системы с распределенными параметрами.

Abstract

Problem of the synthesis of the distributed parameters systems is considered by structural theory methods. Control algorithms and algorithmic structures of systems with feedbacks are determined for the concentrated and distributed control actions.

Введение

Широкий круг управляемых процессов в самых различных областях техники относится к числу сложных динамических объектов с распределенными параметрами (ОРП) [1]. Актуальная задача синтеза систем автоматического управления (САУ) пространственно-распределенными управляемыми величинами ОРП отличается, по сравнению с сосредоточенными системами (ССП), существенной спецификой, обусловленной, во-первых, необходимостью использования в замкнутых контурах регулирования пространственно-зависимых сигналов обратных связей по выходу ОРП в целях формирования соответствующих распределенных или сосредоточенных управляющих воздействий и, во-вторых, трудностями реализации необходимых измерительных и регулирующих устройств.

Пространственная распределенность управляемых величин и обратных связей приводят к принципиальным особенностям, требующим далеко не тривиального развития известных результатов теории ССП в целях разработки способов построения САУ ОРП.

Эффективным аппаратом для решения этой проблемы может служить современная структурная теория распределенных систем, базирующаяся на соответствующих обобщениях понятий объекта управления, передаточной функции динамических звеньев и их соединений [2-4]. В настоящей работе на этой основе предлагаются регулярные методы синтеза алгоритмов управления в САУ ОРП применительно к типовым моделям распределенных объектов [4] и различным видам управляющих воздействий, в том числе, с учетом практических возможностей реализации требуемого характера их пространственного распределения.

Получаемые результаты распространяются на задачи синтеза САУ распределенных объектов сложной структуры.

Весь анализ проводится в детерминированной постановке для идеализированного варианта, предусматривающего возможность осуществления полного и безинерционного измерения текущего состояния вполне наблюдаемого ОРП, хотя в действительности обратные связи по распределенному выходу ОРП могут быть получены лишь с некоторой погрешностью только путем их конструирования с помощью соответствующих наблюдателей по оценкам состояния объекта при его всегда *неполном* измерении [5-8].

Достаточным основанием для такой идеализации (применительно к модальному описанию ОРП [4]) является известная теорема разделения, согласно которой задачу построения системы управления с обратными связями по оценкам состояния можно разделить на две незави-

симые друг от друга задачи синтеза регулятора с обратными связями непосредственно по самим наблюдаемым переменным состояния и проектирования наблюдателя с требуемыми динамическими свойствами, эффективно функционирующего в условиях воздействия помех [5, 6].

1. Алгоритмы и системы управления типовыми ОРП с распределенными управляющими воздействиями

Будем рассматривать типовой ОРП, описываемый относительно управляемой функции его состояния $Q(x, t)$, $x \in [x_0, x_1]$, $t \geq 0$ пространственно одномерным линейным и стационарным уравнением второго порядка в частных производных со стандартными однородными граничными и нулевыми начальными условиями для стандартизирующего распределенного входного воздействия $w(\xi, t)$, $\xi \in [x_0, x_1]$ [4].

Здесь x и ξ - пространственные аргументы $Q(x, t)$ и $w(\xi, t)$ соответственно, и t - время. Передаточная функция $W(x, \xi, p)$ такого объекта от его входа $\tilde{w}(\xi, p)$ к выходу $\tilde{Q}(x, p)$ зависящая от трех аргументов, вместо одного для ССП, представляет собой изображение по временной переменной его функции Грина [2] и может быть получена в форме следующего разложения в бесконечный ряд по собственным функциям $\varphi_n(\mu_n, x)$ краевой задачи, моделирующей поведение ОРП [4]:

$$(1) \quad W(x, \xi, p) = \sum_{n=1}^{\infty} W_{1n}(p) \varphi_n(\mu_n, x) \varphi_n(\mu_n, \xi) r(\xi).$$

Здесь p - оператор Лапласа; μ_n^2 и $r(\xi)$ - собственные числа и весовая функция конечного интегрального преобразования $Q(x, t)$ с ядром $\varphi(\mu, x)$, и

$$(2) \quad W_{1n}(p) = \frac{1}{Ap^2 + A_1p + \mu_n^2},$$

где A и A_1 - постоянные коэффициенты при производных $\partial^2 Q / \partial t^2$ и $\partial Q / \partial t$ соответственно в уравнениях объекта.

Обобщенная структурная схема САУ типового ОРП, построенной по принципу отклонения, приобретает стандартный вид (рисунок 1), однако кардинально отличается при этом от аналогичной структуры для ССП пространственно-распределенным характером информационных сигналов в замкнутом контуре.

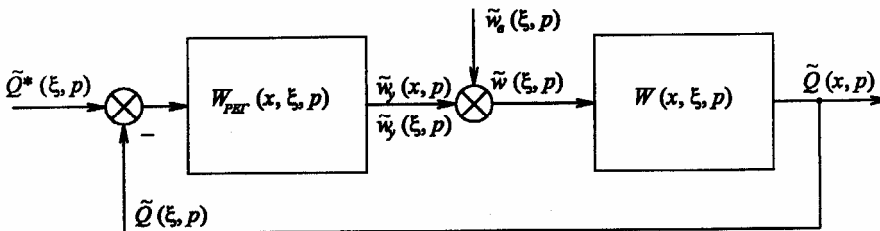


Рисунок 1 – Структурная схема замкнутой системы распределенного управления распределенным выходом ОРП

Регулятор, представляющий собой некоторый распределенный блок [2-4] с передаточной функцией $W_{PEГ}(x, \xi, p)$, учитывающей характеристики датчиков обратной связи и регулирующих органов, формирует распределенное управляющее воздействие $\tilde{w}_y(\xi, p)$ на входе объекта в функции отклонения управляемого распределенного выхода $\tilde{Q}(x, p)$ от его заданного состояния $\tilde{Q}^*(x, p)$. Вторая составляющая $\tilde{w}_e(\xi, p)$ стандартизирующей функции $\tilde{w}(\xi, p)$, учитывающая влияние граничных и начальных условий, моделирует воздействие по внешним возмущениям.

Задача состоит в определении передаточной функции регулятора, обеспечивающей требуемые статические и динамические свойства САУ, и построении структурных схем замкнутой системы, реализующих найденные алгоритмы управления.

Согласно известным правилам структурных преобразований для соединений распределенных блоков [2-4], передаточная функция $W_3(x, \xi, p)$ замкнутой системы (рисунок 1) от задающего воздействия $\tilde{Q}^*(x, p)$ к распределенному выходу $\tilde{Q}(x, p)$ является решением интегрального уравнения:

$$(3) \quad W_3(x, \xi, p) + \int_{x_0}^{x_1} W_{\Pi}(x, v, p) W_3(v, \xi, p) dv = W_{\Pi}(x, \xi, p),$$

$$\text{где } W_{\Pi}(x, \xi, p) = \int_{x_0}^x W(x, \eta, p) W_{PEГ}(\eta, \xi, p) d\eta.$$

Ограничимся далее случаем, когда регулятор описывается рядом вида (1) или (практически всегда) частичной суммой $N < \infty$ его первых членов [3]. Тогда

$$(4) \quad W_{PEГ}(x, \xi, p) = \sum_{n=1}^N W_{2n}(p) \varphi_n(\mu_n, x) \varphi_n(\mu_n, \xi) r(\xi),$$

где $N \leq \infty$, и $W_{2n}(p)$ – некоторые подлежащие определению реализуемые дробно-рациональные функции комплексного аргумента, например, типа (2) с другими коэффициентами.

Используя свойства ортогональности и ортонормированности собственных функций $\varphi_n(\mu_n, x)$ [2-4], можно после подстановки выражений (1) и (4) в (3) и простых преобразований получить точное решение уравнения (3) в явной форме:

$$(5) \quad W_3(x, \xi, p) = \sum_{n=1}^N \frac{W_{1n}(p) W_{2n}(p)}{1 + W_{1n}(p) W_{2n}(p)} \varphi_n(\mu_n, x) \varphi_n(\mu_n, \xi) r(\xi).$$

Поскольку выход $\tilde{Q}(x, p)$ объекта связывается с задающим воздействием $\tilde{Q}^*(\xi, p)$ операцией пространственной композиции [2-4]

$$(6) \quad \tilde{Q}(x, p) = \int_{x_0}^{x_1} W_3(x, \xi, p) \tilde{Q}^*(\xi, p) d\xi,$$

то после подстановки $W_3(x, \xi, p)$ в форме (5), будем иметь:

$$(7) \quad \tilde{Q}(x, p) = \sum_{n=1}^{\infty} W_n(p) \overline{\tilde{Q}}_n^*(\mu_n, p) \varphi_n(\mu_n, x).$$

Здесь

$$(8) \quad W_n(p) = \begin{cases} \frac{W_{1n}(p)W_{2n}(p)}{1+W_{1n}(p)W_{2n}(p)}, & n = \overline{1, N}; \\ 0, & n > N, \end{cases}$$

а $\overline{Q}_n^*(\mu_n, p)$ есть конечное интегральное преобразование (n -я временная мода) $\tilde{Q}^*(\xi, p)$:

$$(9) \quad \overline{Q}_n^*(\mu_n, p) = \int_{x_0}^{x_1} \tilde{Q}^*(\xi, p) \varphi_n(\mu_n, \xi) r(\xi) d\xi,$$

причем $\tilde{Q}^*(\xi, p)$ восстанавливается по известным значениям мод $\overline{Q}_n^*(\mu_n, p)$ в форме бесконечного ряда [3, 4]:

$$(10) \quad \tilde{Q}^*(\xi, p) = \sum_{n=1}^{\infty} \overline{Q}_n^*(\mu_n, p) \varphi_n(\mu_n, \xi).$$

Аналогичное представление $\tilde{Q}(x, p)$ в форме ряда вида (10):

$$(11) \quad \tilde{Q}(x, p) = \sum_{n=1}^{\infty} \overline{Q}_n(\mu_n, p) \varphi_n(\mu_n, x),$$

где моды $\overline{Q}_n(\mu_n, p)$ функции состояния $\tilde{Q}(x, p)$ определяются подобно (9), приводит после сравнения выражений (7) и (11) к бесконечной системе равенств:

$$(12) \quad \overline{Q}_n(\mu_n, p) = W_n(p) \overline{Q}_n^*(\mu_n, p), \quad n = 1, 2, \dots$$

Эти равенства означают, что каждую из N первых временных мод $\overline{Q}_n(\mu_n, p)$ управляемого выхода объекта можно рассматривать в замкнутой системе в качестве выходной величины «элементарного» замкнутого контура регулирования со «своей» передаточной функцией (8), формируемой для «элементарного» объекта $W_{1n}(p)$ «индивидуальным» регулятором $W_{2n}(p)$, и «своим» автономным входным воздействием по величине соответствующей временной моды $\overline{Q}_n^*(\mu_n, p)$ распределенного задания $\tilde{Q}^*(\xi, p)$, независимо от других входов и выходов таких контуров (рисунок 2).

Из равенств (12) также следует, что система не пропускает, в соответствии с (8), моды $\overline{Q}_n^*(\mu_n, p)$ задания $\tilde{Q}^*(\xi, p)$ для $n > N$. Выход объекта не содержит этих мод в сумме ряда (11) при воздействии $\tilde{Q}^*(\xi, p)$, и они оказываются, по существу, нерегулируемыми.

Задающие воздействия и сигналы обратных связей в каждом из независимых контуров регулирования (или сразу разностные сигналы $\overline{Q}_n^*(\mu_n, p) - \overline{Q}_n(\mu_n, p)$ по отклонению $\tilde{Q}^*(\xi, p) - \tilde{Q}(\xi, p)$) должны быть получены по информации о заданном и текущем состояниях объекта с помощью специализированных вычислительных устройств (анализаторов [7]) H_{Q^*}, H_Q (или сразу $H_{\Delta Q}$), представляющих собой совокупность N интеграторов вида (9).

Сумматоры $\Sigma_Q^{(y)}$ и $\Sigma_Q^{(e)}$ восстанавливают распределенный выход объекта в форме ряда (11).

Сигналы на выходах регуляторов $W_{2n}(p)$ (рисунок 2) представляют собой моды $\overline{w}_{yn}(\mu_n, p)$ распределенного управляющего воздействия $\tilde{w}_y(\xi, p)$ [4], по которым оно должно

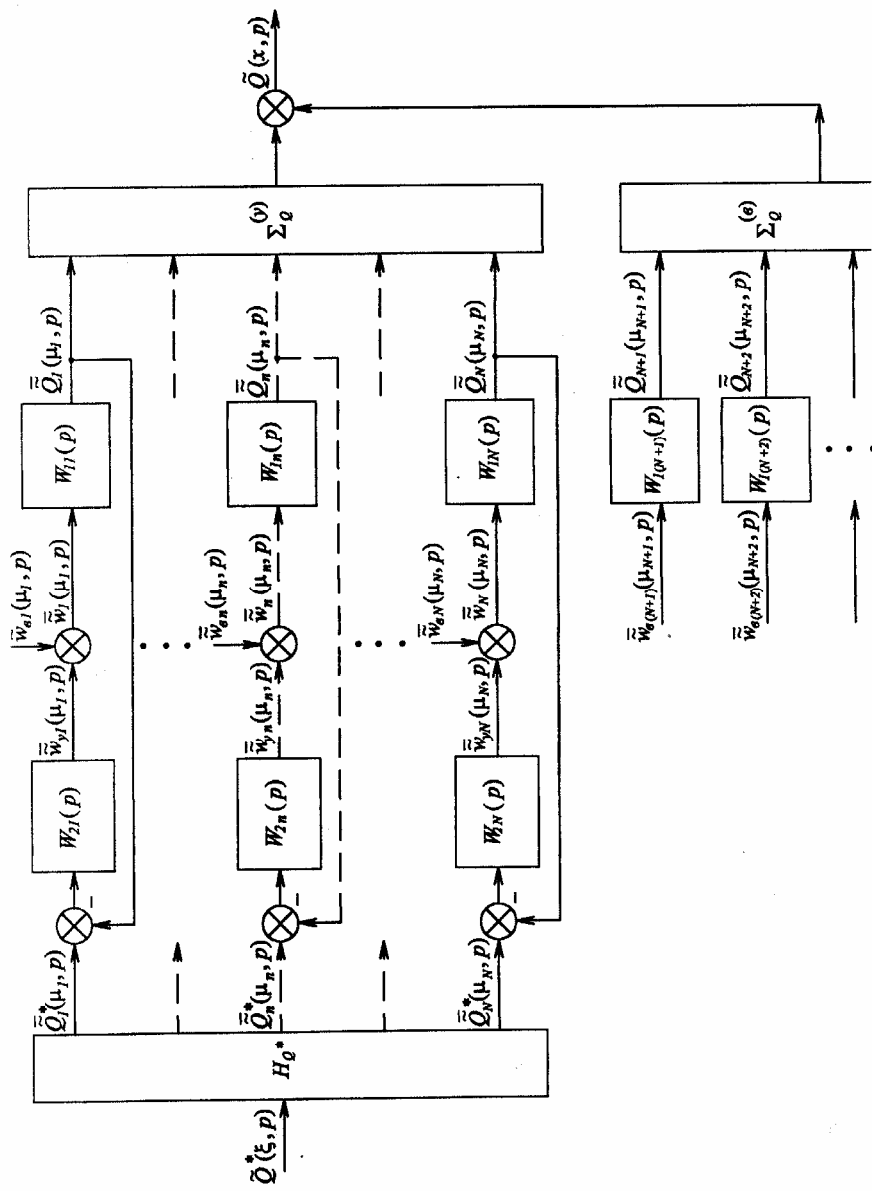


Рисунок 2 – Структурная схема системы управления распределенным выходом ОРП с независимым регулированием временных мод функции состояния объекта

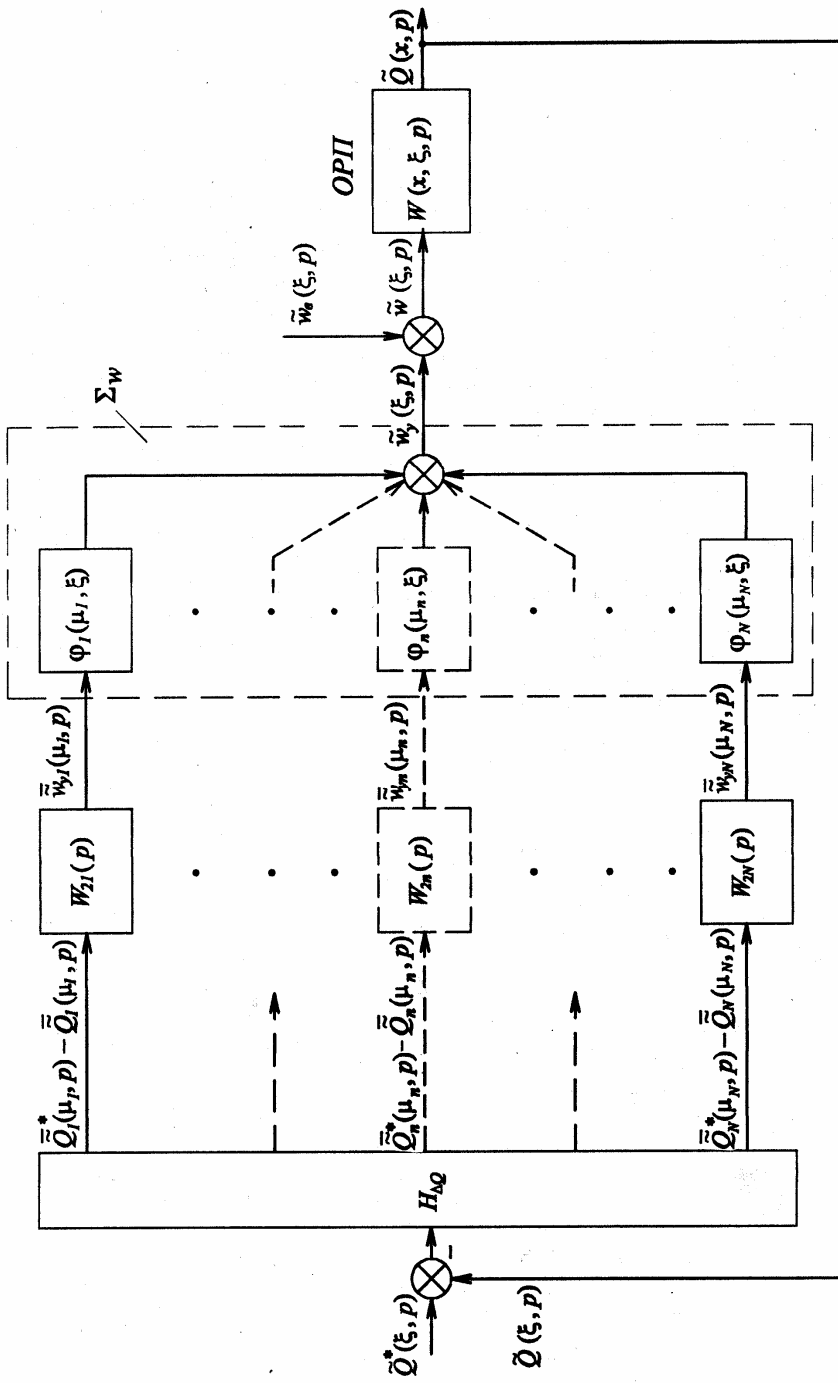


Рисунок 3 – Вариант структурной схемы системы управления с анализатором отклонений управляемой величины от ее заданного состояния

быть восстановлено с весовыми множителями $\varphi_n(\mu_n, \xi)$ на выходе специального сумматора \sum_w (синтезатора [7]), работающего по алгоритму, подобному (10) и (11):

$$(13) \quad \bar{w}_y(\xi, p) = \sum_{n=1}^N \bar{w}_{yn}(\mu_n, p) \varphi_n(\mu_n, \xi).$$

Структурная схема реализуемой САУ с анализатором $H_{\Delta Q}$ и синтезатором \sum_w показана на рисунке 3.

Выбор регуляторов $W_{2n}(p)$ в каждом из контуров регулирования мод выхода объекта можно проводить по обычным методикам синтеза замкнутых ССП из условий обеспечения допустимой статической точности при воздействии возмущений (в виде мод $\bar{w}_{en}(\mu_n, p)$ составляющей $\bar{w}_e(\xi, p)$ стандартизирующего входа $\bar{w}(\xi, p)$ (рисунки 1, 2)) и желаемого поведения в динамических режимах работы. Усложняющим обстоятельством здесь является необходимость определения требуемых показателей качества с учетом характеристик неуправляемых составляющих $\tilde{Q}(x, p)$, которые создают некомпенсируемые системой (в условиях $W_n(p) = 0, n > N$ в (8)) отклонения от задания $Q^*(\xi, t)$, уменьшающиеся с ростом N . Последнее обстоятельство может служить основанием для выбора числа N управляемых мод объекта.

Устойчивость системы управления ОРП (1) с регулятором (4) в целом обеспечивается, с одной стороны, лишь при наличии достаточного запаса устойчивости в каждом из автономно управляемых контуров, что, в свою очередь, выполняется в условиях отсутствия неограниченного приближения к границе области устойчивости при $N \rightarrow \infty$ [2], а, с другой стороны, при сохранении устойчивости разомкнутых контуров регулирования всех остальных мод для $n > N$, если $N < \infty$.

2. Алгоритмы и системы управления типовыми ОРП с фиксированным характером пространственного распределения управляющих воздействий

Поскольку закономерности пространственного распределения управляющих воздействий часто диктуются конкретными возможностями их реализации на объектах управления вполне определенного конструктивного исполнения, то типичной при практическом построении САУ ОРП является ситуация, когда эти закономерности оказываются заранее заданными соответствующими проектными решениями объекта и не совпадают с требуемыми зависимостями вида (13). В таком случае распределенные управляющие воздействия заведомо фиксируются, как правило, в форме

$$(14) \quad w_y(\xi, t) = \sum_{m=1}^s v_{ym}(t) \varphi_{ym}(\xi)$$

с известными распределениями $\varphi_{ym}(\xi)$, $m = \overline{1, s}$, $s \geq 1$, чаще всего, простейшего вида, например, заданными в классе кусочно-постоянных функций или локальных воздействий, моделируемых дельта-функциями [7]. В этой ситуации задача сводится к синтезу САУ с s сосредоточенными управлениями $v_{ym}(t)$, $m = \overline{1, s}$, в (14) в условиях сохранения управляемости ОРП при подобных ограничениях. Для формирования каждого воздействия $\bar{v}_{ym}(p)$ требуется «свой» регулятор $W_{mPEГ}(x, \xi, p)$ с передаточной функцией

$$(15) \quad W_{mPEГ}(x, \xi, p) = W_{\xi m}^*(\xi, p) \varphi_{ym}(x), \quad m = \overline{1, s}.$$

Здесь распределенный ξ -блок $W_{\xi m}^*(\xi, p)$ [2, 4] формирует на выходе каждого m -го регулятора сосредоточенный сигнал $\tilde{v}_{ym}(p)$, описываемый пространственной композицией вида (6) [4]:

$$(16) \quad \tilde{v}_{ym}(p) = \int_{x_0}^{x_1} W_{\xi m}^*(\xi, p) [\tilde{Q}^*(\xi, p) - \tilde{Q}(\xi, p)] d\xi.$$

В итоге распределенный регулятор в схеме рисунка 1 должен иметь передаточную функцию

$$(17) \quad W_{PEГ}(x, \xi, p) = \sum_{m=1}^s W_{\xi m}^*(\xi, p) \varphi_{ym}(x),$$

и тогда задача сводится к определению всех $W_{\xi m}^*(\xi, p)$, $m = \overline{1, s}$. Выберем эти блоки так, чтобы выражение (17) приняло вид (4) для $N = s$:

$$(18) \quad W_{PEГ}(x, \xi, p) = \sum_{n=1}^s W_{n PEГ}^*(p) \varphi_n(\mu_n, x) \varphi_n(\mu_n, \xi) r(\xi),$$

где $W_{n PEГ}^*(p)$ выполняют роль регуляторов $W_{2n}(p)$ в (4).

Если представить все $\varphi_{ym}(x)$ разложениями в бесконечные ряды по собственным функциям $\varphi_n(\mu_n, x)$ с коэффициентами $\overline{\varphi}_{ym}(\mu_n)$, подобно (10)-(11), и ограничиться в первом приближении учетом только s первых членов этих рядов, то равенство сумм в (17) и (18) сводится к выполнению s равенств

$$(19) \quad \sum_{m=1}^s \overline{\varphi}_{ym}(\mu_n) W_{\xi m}^*(\xi, p) = W_{n PEГ}^*(p) \varphi_n(\mu_n, \xi) r(\xi), \quad n = \overline{1, s},$$

которые можно рассматривать как систему s линейных уравнений относительно неизвестных $W_{\xi m}^*(\xi, p)$ при заданных передаточных функциях $W_{n PEГ}^*(p)$.

При неособенной квадратной матрице Φ системы (19)

$$(20) \quad \Phi = [\overline{\varphi}_{ym}(\mu_n)]; \quad m, n = \overline{1, s}$$

для объекта, управляемого по заданному состоянию, ее решение определяется формулами Крамера в форме линейной комбинации:

$$(21) \quad W_{\xi m}^*(\xi, p) = \sum_{n=1}^s \frac{D_{nm}^*}{D} W_{n PEГ}^*(p) \varphi_n(\mu_n, \xi) r(\xi), \quad m = \overline{1, s},$$

где D – определитель матрицы Φ , и D_{nm}^* – алгебраическое дополнение n -го элемента m -го столбца D .

Теперь, согласно выводам п.1, система управления ОРП должна строиться в виде совокупности независимых контуров регулирования s мод объекта $\tilde{Q}_n(\mu_n, p)$, $n = \overline{1, s}$, с помощью автономных регуляторов $W_{n PEГ}^*(p)$, $n = \overline{1, s}$, и остается лишь найти алгоритм формирования действительного управляющего воздействия (14) по выходным сигналам этих регуляторов, используя соотношения (21).

Каждое управление $\tilde{v}_{ym}(p)$ на выходе блока $W_{\xi m}^*(\xi, p)$ описывается выражением (16), которое после подстановок (10), (11), (21) и интегрирования с учетом ортонормированности собственных функций принимает следующий вид:

$$(22) \quad \tilde{v}_{ym}(p) = \sum_{n=1}^s \frac{D_{nm}^*}{D} W_{nPEГ}^*(p) \left[\bar{Q}_n^*(\mu_n, p) - \bar{Q}_n(\mu_n, p) \right], \quad m = \overline{1, s}.$$

Поскольку выходные сигналы $\bar{w}_{yn}^*(\mu_n, p)$ регуляторов $W_{nPEГ}^*(p)$ связываются с сигналами на их входах соотношениями (рисунок 2):

$$(23) \quad \bar{w}_{yn}^*(\mu_n, p) = W_{nPEГ}^*(p) \left[\bar{Q}_n^*(\mu_n, p) - \bar{Q}_n(\mu_n, p) \right], \quad n = \overline{1, s},$$

то вместо (22) будем иметь:

$$(24) \quad \tilde{v}_{ym}(p) = \sum_{n=1}^s \frac{D_{nm}^*}{D} \bar{w}_{yn}^*(\mu_n, p), \quad m = \overline{1, s},$$

и, следовательно, распределенное управление (14) определяется суммой

$$(25) \quad \tilde{w}_y(\xi, p) = \sum_{m=1}^s \left[\sum_{n=1}^s \frac{D_{nm}^*}{D} \bar{w}_{yn}^*(\mu_n, p) \right] \varphi_{ym}(x).$$

В итоге структурная схема САУ отличается от схемы рисунка 3 лишь устройством синтезатора Σ_w^* , который выполняет операцию взвешенного суммирования (24) выходов автономных регуляторов и формирует заведомо реализуемое управляющее воздействие на входе ОРП согласно алгоритму (25).

В простейшем частном случае $s=1$, широко распространенном на практике, можно выбрать единственный регулятор $W_{PEГ}(x, \xi, p)$ вида (15) с ξ -блоком $W_{\xi}^*(\xi, p)$ в форме, подобной (4):

$$(26) \quad W_{\xi}^*(\xi, p) = W_{PEГ}^*(p) \sum_{m=1}^N k_m \varphi_m(\mu_m, \xi) r(\xi),$$

где k_m – некоторые весовые множители, и $W_{PEГ}^*(p)$ – подлежащая определению дробно-рациональная передаточная функция, характеризующая статические и динамические свойства регулятора (26). Анализ по схеме п.1 приводит здесь к равенству

$$(27) \quad \sum_{n=1}^N k_n \bar{Q}_n(\mu_n, p) = \frac{W_{PEГ}^*(p) \sum_{n=1}^N k_n W_{1n}(p) \bar{\varphi}_y(\mu_n)}{1 + W_{PEГ}^*(p) \sum_{m=1}^N k_m W_{1m}(p) \bar{\varphi}_y(\mu_m)} \sum_{m=1}^N k_m \bar{Q}_m^*(\mu_m, p),$$

где $\bar{\varphi}_y(\mu_n)$ – моды разложения в бесконечный ряд по системе $\varphi_n(\mu_n, x)$ функции $\varphi_y(x) = \varphi_{y1}(x)$ при $s=1$ в (14).

Если принять здесь $k_n, n = \overline{1, N}$, равными значениям собственных функций в некоторой точке $x^0 \in [x_0, x_1]$:

$$(28) \quad k_n = \varphi_n(\mu_n, x^0), \quad n = \overline{1, N},$$

то при $N \rightarrow \infty$ равенство (27) приобретает следующий вид:

$$(29) \quad \tilde{Q}(x^0, p) = \frac{W_{PEГ}^*(p) W_x^{(y)}(x^0, p)}{1 + W_{PEГ}^*(p) W_x^{(y)}(x^0, p)} \tilde{Q}^*(x^0, p).$$

Здесь

$$(30) \quad W_x^{(y)}(x^0, p) = \sum_{n=1}^{\infty} W_{1n}(p) \bar{\varphi}_y(\mu_n) \varphi_n(\mu_n, x^0)$$

есть передаточная функция от сосредоточенного управления $\tilde{v}_y(p)$ к выходу объекта в точке $x = x^0$ [4], представляющего собой в таком случае распределенный x -блок [2-4].

При $W_{PEГ}^*(p) \gg 1$ в силу (29) с требуемой точностью выполняется равенство

$$(31) \quad \tilde{Q}(x^0, p) \cong \tilde{Q}^*(x^0, p),$$

и, следовательно, на выходе системы с погрешностью, определяемой выбором конечного числа N и параметров регулятора $W_{PEГ}^*(p)$ устанавливается заданное значение функции состояния ОРП в выбранной точке $x = x^0$.

При этом пространственное распределение выхода ОРП однозначно фиксируется его свойствами независимо от управляющих воздействий регулятора, согласно соотношению

$$(32) \quad \tilde{Q}(x, p) = \frac{W_x^{(y)}(x, p)}{W_x^{(y)}(x^0, p)} \tilde{Q}(x^0, p).$$

В частном случае локального характера воздействия $\tilde{v}_y(p)$ в точке $x_1^* \in [x_0, x_1]$ распределение $\varphi_y(x)$ представляет собой передаточную функцию переходного x -блока [2-4]:

$$(33) \quad \varphi_y(x) = \psi_1(x) \delta(x - x_1^*) + \psi_2(x) \delta'(x - x_1^*), \quad \psi_1^2 + \psi_2^2 > 0.$$

Здесь $\delta(x - x_1^*)$ и $\delta'(x - x_1^*)$ – пространственная дельта-функция, сосредоточенная в точке $x = x_1^*$ и ее первая производная, а $\psi_1(x)$ и $\psi_2(x)$ – множители, определяемые конкретным характером краевой задачи, описывающей поведение объекта, и местом приложения локального воздействия [4].

Если x_1^* совпадает с одной из граничных точек x_0 или x_1 пространственной области, занимаемой объектом, то мы получаем наиболее просто реализуемую систему *граничного управления*. В этом случае для типичного варианта $\psi_1 = const$, $\psi_2 = 0$ [4] будем иметь в (27)

$$(34) \quad \bar{\varphi}_y(\mu_n) = \psi_1 \varphi_n(\mu_n, x_1^*) r(x_1^*).$$

Число N регулируемых мод и параметры регулятора $W_{PEГ}^*(p)$ при локальном сосредоточенном воздействии могут быть предварительно найдены, исходя из требуемой точности поддержания заданного состояния ОРП (например, в установившемся режиме работы САУ) с учетом воздействия возмущений и на нерегулируемые моды.

Передаточная функция разомкнутой системы в соответствии с выражением (27) при выборе k_n , согласно (28), записывается в виде

$$(35) \quad W_{раз}(p) = W_{PEГ}^*(p) \sum_{m=1}^N W_{1m}(p) \bar{\varphi}_y(\mu_m) \varphi_m(\mu_m, x^0).$$

Если при граничном управлении в условиях (34) принять здесь $x^0 = x_1^*$, то передаточная функция (35) имеет заведомо положительные коэффициенты при всех $W_{1m}(p)$:

$$(36) \quad W_{раз}(p) = W_{PEГ}^*(p) \sum_{m=1}^N W_{1m}(p) \varphi_m^2(\mu_m, x_1^*) r(x_1^*).$$

В таком случае для характерного варианта $A=0$ в (2) заведомо обеспечивается устойчивость замкнутой системы при любом коэффициенте передачи $k_{PEГ}^*$ пропорционального регулятора $W_{PEГ}^*(p) = k_{PEГ}^*$.

По данным передаточным функциям с помощью типовых частотных методов можно окончательно выбрать регулятор $W_{PEГ}^*(p)$, обеспечивающий желаемый запас устойчивости и необходимые показатели качества переходных процессов для всех управляемых мод. Устойчивость всей системы в целом, как и для распределенных управлений, обеспечивается только в случае устойчивости разомкнутых контуров регулирования всех остальных мод объекта.

3. Алгоритмы и системы управления распределенными объектами сложной структуры

Предположим, что передаточная функция ОРП достаточно сложной структуры, описываемого в общем случае некоторой системой линейных уравнений в частных производных [1, 2, 4], представима в подобной (1) форме разложения в сходящийся в бесконечный ряд:

$$(37) \quad W(x, \xi, p) = \sum_{n=1}^{\infty} \hat{W}_{1n}(p) \hat{\varphi}_n(x) \hat{\varphi}_n(\xi) \hat{r}(\xi),$$

где $\hat{W}_{1n}(p)$ отличаются, вообще говоря, от передаточных функций (2), а $\hat{\varphi}_n(x)$, $n=1,2, \dots$ – некоторое семейство ортогональных и ортонормированных с весом $\hat{r}(x)$ на отрезке $[x_0, x_1]$ функций, *не обязательно совпадающих с собственными функциями объекта*. Если такое описание сложного ОРП найдено, то задача синтеза распределенного регулятора вида (4)

$$(38) \quad W_{PEГ}(x, \xi, p) = \sum_{n=1}^N \hat{W}_{2n}(p) \hat{\varphi}_n(x) \hat{\varphi}_n(\xi) \hat{r}(\xi)$$

решается по схеме, предложенной в п.1, поскольку все полученные выше в п.1 выводы используют применительно к собственным функциям $\varphi_n(\mu_n, x)$ только свойство их ортогональности и ортонормированности, сохраняющееся для семейства $\hat{\varphi}_n(x)$, и эти выводы не зависят от конкретного вида $W_{1n}(p)$. В итоге, САУ объектами (37) опять строится по схемам рисунков 1-3 как совокупность N независимых контуров управления звеньями $\hat{W}_{1n}(p)$ с помощью автономных регуляторов $\hat{W}_{2n}(p)$, $n = \overline{1, N}$, если предположить, что функция состояния объекта $\tilde{Q}(x, p)$ (включая заданную величину $\tilde{Q}^*(x, p)$) может быть представлена ее разложением в бесконечный ряд по системе $\hat{\varphi}_n(x)$, аналогично (9)-(11):

$$(39) \quad \tilde{Q}(x, p) = \sum_{n=1}^{\infty} \overline{\tilde{Q}}_n(p) \hat{\varphi}_n(x), \quad \overline{\tilde{Q}}_n(p) = \int_{x_0}^{x_1} \tilde{Q}(x, p) \hat{\varphi}_n(x) r(x) dx,$$

и подобным образом будем иметь для входного стандартизирующего воздействия

$$(40) \quad \tilde{w}(\xi, p) = \sum_{n=1}^{\infty} \overline{\tilde{w}}_n(p) \hat{\varphi}_n(\xi); \quad \overline{\tilde{w}}_n(p) = \int_{x_0}^{x_1} \tilde{w}(\xi, p) \hat{\varphi}_n(\xi) r(\xi) d\xi.$$

Аналогичным путем на модели ОРП вида (37) распространяются результаты, полученные в п. 2. Таким образом, проблема сводится к представлению передаточной функции объекта сложной структуры в «типовой» форме (37).

Если теперь $W(x, \xi, p)$ записывается в форме (37), то, операция пространственной композиции для вход-выходного соотношения ОРП [2-4] приводится к виду:

$$(41) \quad \tilde{Q}(x, p) = \int_{x_0}^{x_1} W(x, \xi, p) \tilde{w}(\xi, p) d\xi = \sum_{n=1}^{\infty} \hat{W}_{1n}(p) \bar{w}_n(p) \hat{\phi}_n(x).$$

Сравнивая (39) и (41), находим совпадающее с характеристиками типового ОРП (рисунок 2) соотношение:

$$(42) \quad \bar{Q}_n(p) = \hat{W}_{1n}(p) \bar{w}_n(p), \quad n = 1, 2, \dots,$$

согласно которому $\hat{W}_{1n}(p)$ представляет собой передаточную функцию

$$(43) \quad \hat{W}_{1n}(p) = \frac{\bar{Q}_n(p)}{\bar{w}_n(p)}, \quad n = 1, 2, \dots$$

по n -ым составляющим рядов (39) и (40).

Согласно принципу суперпозиции для линейных объектов, каждая из составляющих $\bar{Q}_n(p) \hat{\phi}_n(x)$ ряда (39) описывается решением уравнений ОРП как реакция на соответствующее слагаемое $\bar{w}_n(p) \hat{\phi}_n(\xi)$ входного воздействия (40) в силу равенства (43). Такие решения при соответствующем выборе семейства $\hat{\phi}_n(x)$, как правило, могут быть найдены известными методами [2-4, 9]. Тогда по ним находятся передаточные функции (47), и, в соответствии с соотношениями (39)-(42), объект в целом оказывается представимым в форме (37) при выполнении указанных выше практически малостепенных допущений. В более сложных ситуациях с пространственно-многомерным выходом объекта описанный подход к синтезу САУ может быть реализован путем представления $W(x, \xi, p)$, $W_{\text{рег}}(x, \xi, p)$, $Q(x, p)$ и $w(\xi, p)$ в форме их разложений в кратные ряды по соответствующим системам ортогональных функций.

К аналогичным результатам приводит метод пространственно-частотной декомпозиции при синтезе САУ температурным полем в многослойной среде, описываемым системой уравнений теплопроводности [10].

Список литературы

- [1] Бутковский А.Г. Методы управления системами с распределенными параметрами. - М.: Наука, 1975.
- [2] Бутковский А.Г. Структурная теория распределенных систем. - М.: Наука, 1977.
- [3] Мартыненко Н.А., Пустыльников Л.М. Конечные интегральные преобразования и их применение к исследованию систем с распределенными параметрами. - М.: Наука, 1986.
- [4] Рапопорт Э.Я. Структурное моделирование объектов и систем управления с распределенными параметрами. - М.: Высшая школа, 2003.
- [5] Андреев Ю.Н. Управление конечномерными объектами. - М.: Наука, 1976.
- [6] Первозванский А.А. Курс теории автоматического управления. - М.: Наука, 1986.
- [7] Рей У. Методы управления технологическими процессами. - М.: Мир, 1983.
- [8] Дегтярев Г.Л., Сиразетдинов Т.К. Теоретические основы оптимального управления упругими космическими аппаратами. - М.: Машиностроение, 1986.
- [9] Карташов Э.М. Аналитические методы в теории теплопроводности твердых тел. - М.: Высшая школа, 2001.
- [10] Коваль В.А. Спектральный метод анализа и синтеза распределенных управляемых систем. - Саратов: СГТУ, 1997.

EFFECTIVE MANAGEMENT TECHNIQUES FOR INTERNATIONAL ENGINEERING PROJECTS

R.H. Rigby¹, V.A. Zarubin²

¹Fluor Daniel, Inc. One Fluor Daniel Drive, Sugar Land, Texas 77478, USA
Hugh.Rigby@Fluor.com, hugh.rigby@att.net
phone: +1-281-263-1000

²Institute for Control of Complex Systems, Russian Academy of Sciences
Sadovaya, 61, 443020, Samara, Russia
Zarubin@iccs.ru, zarubin@shaw.ca
phone: +7-8462-323927

Keywords: *open system, management, project, process, goals, sensitivity analysis, SWOT*

Abstract

В работе обсуждается использование некоторых эффективных методов управления для решения проблем, которые могут возникнуть в международных технических проектах. Сначала описываются проблемы, затем исследуется ряд возможных действий по их решению. Создается гипотетическая проектная модель «как есть». На проектную систему направляется влияние четырех различных факторов и анализируется реакция системы на эти воздействия. Оцениваются сильные и слабые стороны произведенных изменений с использованием аналитического метода SWOT. Предложена реорганизация бизнес-процессов с использованием результатов, полученных при помощи SWOT-анализа, и создана модель «как должно быть».

Introduction

A hypothetical international engineering project is taken into consideration [1]. The project has a complex organization structure and contains multilevel management systems that include:

- Owner's engineering coordination and project ownership office
- Project management office including a central project engineering management
- Local engineering and project management offices
- Design institutes and engineering companies offices
- Construction contractors offices and on-site field support
- Construction subcontractors offices and contractor engineering support
- Local, regional and national governmental regulation (certification) offices

These organizations are geographically dispersed and can often be located in different countries and separated by great distances. Making the situation even more complex is the fact that all participants involved in the project have their own hierarchical organizational structures. Therefore, collaboration in mutual engineering and construction processes is refracted via prism of management structures within each organization. As a result, there is the potential for a number of issues that can affect the project execution. In order not to be too lengthy, this paper discusses only a few of these issues that can impact a project, as shown in Section 2. Yet this approach should convey some very important ideas. The initial state of the project is fixed in the AS IS model. A sensitivity analysis of the AS IS model allows understanding system reactions to the introduction of different technological and organizational impacts (changes) that need to be managed. This sensitivity analysis creates a basis for application of BPR and the creation of a TO BE model. Discipline specialization, creation of independent engineering teams (holons [2]) is an effective way for project efficiency improvement. Integrated multidisciplinary models, integrated databases and concurrent engineering form a basis for ef-

fective engineering execution. Concurrent engineering is the earliest possible simultaneous work of experts from various disciplines using set of multidisciplinary physical models, in order to achieve high quality, functionality and manufacturability in the shortest time, for a minimum cost. Concurrent engineering is primarily an expression of the desire to increase the competitiveness by decreasing the product lead-time, while improving quality and cost.

Decentralization of decision-making happens as a result of introducing these techniques. A corporate information system (CIS) based on the process approach to management [3] has to be designed for the project in accordance with all of these project execution characteristics. These topics are developed within the body of this paper and a conclusion is made at the end. So, here is a closer look at these issues.

1 Problem description and goal of this paper

The issues for the hypothetical project are as follows:

- Dissociation, sluggishness and doubling of functions in hierarchical structures and among participants lead to inefficiency, time increase and cost.
- Production at the lower levels of the hierarchies depends upon the pertinent department heads (managers) and form “castes” that establish information channels from one function to another and religiously blockade all unsanctioned communication or data exchange. The data in such situations is transferring via interdepartmental paths, which contain barriers and fences, and experience further obstacles on higher levels. These barriers from one department to another and from one level to another are inefficient and thus increase time and costs.
- Interaction with governmental structures in different countries is a complicated problem that is often better known to the managers of local engineering and construction companies. These managers together with local governmental authorities can form another “caste” that may pursue their own goals that are different from project’s ones.
- Communications and data exchange is very complicated in such projects. When engineers and technical specialists from different departments, from different companies, from different countries and from different cultures are cooperating in design and construction of one system or object, effective data exchange information system is vital.

Engineering in the AS IS model leads to many design errors that have to be corrected at the construction stage, because unfortunately they have not been recognized until this stage. The cost-time diagram, Figure 1, shows this.

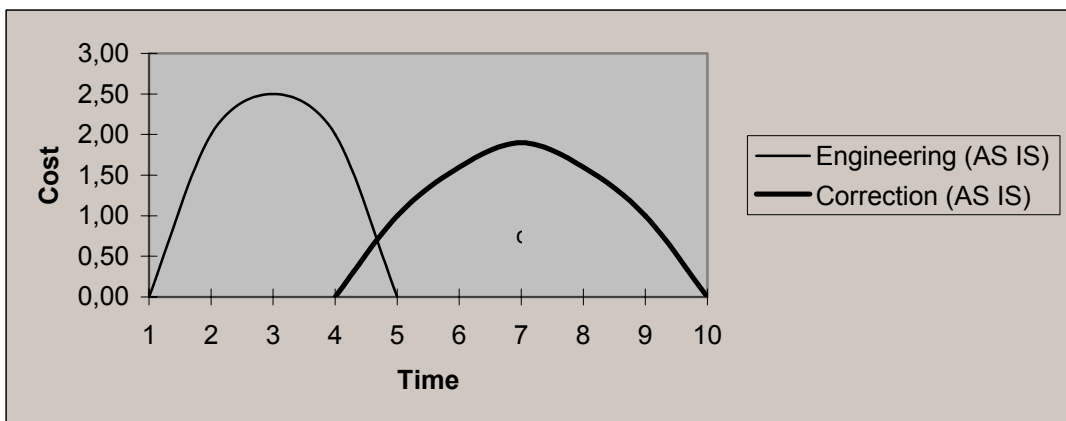


Figure 1 – Relative cost-time relationship for AS IS mode

In this example, detailed engineering works are finished after five months; construction begins after on four months. However, in the AS IS model, engineering works were continued during a project correction period and many discrepancies in engineering are discovered during construction. The cost of correcting such discrepancies is very high due to the mobilization of expensive construction resources, experience rework and delays awaiting resolution of engineering issues.

This paper seeks to analyze the issues described, and to offer a process (BPR) for determining needed changes and the steps to achieving implementation of BPR. Likewise, the process used to accomplish this can be posed as an introduction of a process approach to international engineering management practice. Also, this process will allow the solutions of problems such as those mentioned above and will thus increase the efficiency of international engineering project execution.

To achieve this goal, we have to solve the following tasks:

- Accomplish a sensitivity analysis of the AS IS model for the changes that will occur after introduction of new technologies and methods into the project execution.
- Implement the re-engineering of business-processes (BPR) to achieve the best practice and to creation the TO BE model.
- Make conclusions of the findings and plan further activities.

2 Sensitivity analysis of AS IS model

The AS IS model is created to describe a hypothetical model with potentially realistic situations that can occur within international engineering projects. The model is created with realistic situations so that a useful analysis can be done. Let us fix the situation like it was described in the introduction and provide several possible technological impacts to this state. The reaction to the impacts is the subject of our analysis. The following impacts will be presented: (1) introduction of extranet into management practice, (2) utilizing concurrent engineering methods and practice and integrated project execution (3) flattening of organizational structure (4) decentralization of decision-making within framework of project goals and overall criteria established by owners, creditors and top project management.

2.1 Impact of extranet being used by management

A project web site is set up and organized with extranet standards, creating a good means for effective data exchange such as project goals, scopes definitions, drawings, specifications, standards, design basis, project execution plan, organization charts, announcements, online meetings, videoconferences, schedules, budgets, progress reports and more. This technique allows putting into practice virtual (distributed via the internet) engineering and combining geographically dispersed offices into one virtual office. This approach gives an opportunity for the orchestration and coordination of the various engineering discipline teams that work in different offices. This “virtual office” operates almost as a real “face-to-face” office, yet enables people from different geographical locations to work together. The strong and weak sides of this approach, opportunities and threats are presented in table 1.

Table 1 – SWOT table for introduction of the extranet:

<p>Strong points: Direct links to project data, fast access to the changes made in other disciplines, possibility to communicate directly with all participants from the various disciplines, seeing many things in writing without chance of misinterpretations via conversations yet allowing conferences and discussions when appropriate</p>	<p>Weak points: Possible complex structure of access authority, need for org chart flattening and decentralization of decision-making process, reluctance by individuals to participate and communicate via internet can slow progress, possibility of playing political games and throwing responsibilities “over the fence” to someone else</p>
--	---

Opportunities: Reduction in cost of expatriots living in-country where work may be taking place at offshore engineering offices, wider participation of quality due to eliminating inconveniences of extended travel and being away from home, reduction of travel costs	Threats: Possibility of control lost due to direct horizontal links, technical malfunctions can disrupt progress, vulnerability to in-country governments that may control internet and telecommunication links
--	---

2.2 Impact of using an integrated design model of facilities - drawings and engineering “Deliverables”

Authors stressed in Globalization Problems in International Engineering Project Management, Engineering and Management Journal, Kazakhstan, 2002 [1] that in order to work globally, it is very important to have an integrated model. It’s obvious that effective management needs reliable information about the current state of a project. For geographically dispersed projects, where many managerial groups, design institutes, engineering companies, vendors, construction contractors and subcontractors take part, an integrated project model has to be created combining the total scope of the project and all known project design considerations. It should encompass all considerations such as geographical, environmental, governments, logistics, communication, process, power supply, owners, developers, commercial, engineering disciplines, main equipment (mechanical, electrical, control systems, telecommunications, architectural), land requirements, socio-political, socio-economic, cost, contracting plan, project organization structure, coordination procedures, work process diagrams, and more. Much of the technical portion can be defined in a design basis, which can be a significant portion of the model, itself. And, the model should become a “working document” in that as the project progress, the model evolves and becomes more perfect. The drawings and technical documents needed for permitting and construction certainly move toward a detailed and final state. (Note that the discussion of a model or any of the main topics mentioned in this paper could be a part of an entirely separate presentation, so this discussion is abbreviated.)

Such a model has to appear at earliest design stage and be developed with the project staff and cover through commissioning, start-up and turnover. An integrated model and an integrated database provide a united vision of the designed facilities and allow working in parallel within the various project groups and sub-groups. All project participants can make their own input and have access to inputs made by others.

One thing that is taking place within the model is that all project participants are working within consistent parameters of document format and software, so that as drawings and technical documents develop, there is a “one project” look to the project. But this is not just an appearance. It is a consistency that translates into clear communication and understanding. It also makes the “deliverables” more user friendly and thus more effectively useful, thus improving construction and other usage efficiency. This translates into lower overall project cost.

Today, there exists all the means needed for creation of integrated models, and implementation yields the possibility of having quick and effective access to databases and shared information.

Table 2 – SWOT table for the introduction of an integrated model:

Strong points: Clear understanding of overall project, momentary check of new impacts from one discipline by other disciplines, direct links to project data by all participants, fast access to the changes made in other disciplines, confirmed sharing of knowledge	Weak points: Complex structure of the model, need accurate project profile of standards definition, need for access authority, access needs to have an effective definition for all participants, need for the flattening of the organization structure and allowance for all to use the model, need to have management of the model
--	--

Opportunities: Reduction in cost of interdisciplinary coordination, overall project schedule reduction, project quality improvement, clarity and proof of engineering progress	Threats: Possibility of departmental concerns for loss of control due to direct horizontal links in some of the engineering organizations
--	---

2.3 Impact of flattening the organizational structure

In general, flattening of organizational structure leads to a reduction in the direct and even the indirect cost of management. To increase efficiency of management, efficient methods and work processes have to be implemented, for example those mentioned in subsections 2.1 and 2.2 above. Not only methods such as these are good, but also the hierarchical management system has to be formed as a holistic one [4]. To help achieve this, holons can be created which are self-organized, self-sufficient groups of specialists that can solve posed problems. All needed authority and responsibility have to be delegated to these groups. As long as project groups stay within project design bases and guidelines, decision-making processes can be decentralized in the project and the solution of only strategic problems could be left for centralized management. To put this into practice, BPR has to be implemented and an up-to-date computer system with united and integrated information has to be applied. In other words, the implementation of this change, can be done if extranet usage and project modeling is done as described in subsections 2.1 and 2.2 of this paper. One last consideration for flattening or decreasing the levels of management within the project organization is that sound, skilled and experienced people should be chosen within the project management and engineering organization. Ultimately, the project will only be as successful as the people within the project organization. Skill and experience can often make up for deficiencies elsewhere, but this is not to diminish the need for good project methods and work processes.

Table 3 – SWOT table for the impact of flattening the organizational structure:

Strong points: Distribution of the discipline tasks between the most effective discipline teams, organizing the work process on the basis of supply / demand, high professional solutions reached	Weak points: The need for and increase in interdisciplinary communication, the need for special technology (extranet, integrated model, integrated database) to be implemented
Opportunities: Reduction in cost of decision-making, reduction in overall schedule due to quicker decision-making, reduction in cost of project management personnel	Threats: Risk of control loss due to absence of hierarchy in management structure

2.4 Impact of decentralization of decision-making

The AS IS model performs decision-making in a central engineering and project management office, which segregated geographically from the offices where the detailed engineering is actually taking place. On a global project, these detailed engineering offices are often dispersed across cities, regions and even countries. In many cases, these dispersed engineering offices are capable of solving engineering and technical problems, but are not empowered to do so and such is the case in the AS IS model. These dispersed engineering offices and institutes are also comprised of multi-disciplines and specialists that are experienced in making significant decisions as well. Almost all engineering solutions involve or impact other disciplines, which is the nature of integrated design. These engineering and technical organizations, though set apart from a group of managers at a central decision-making office, have the resources to grasp the issues, understand the problem, offer viable solutions, and implement the needed actions. The SWOT analysis shown below considers the possibility of decision-making process being decentralized and authority delegated from a central engineering office to the appropriate engineering teams best equipped to resolve the pertinent problems.

Table 4 – SWOT table for the impact of decentralized decision-making:

Strong points: Distribution of the decision-making processes between the most effective discipline teams, greater selection of resources for matching the problem with experienced specialists	Weak points: The need for an increases efficiency in interdisciplinary communication - special technology may be required for communication and coordination (extranet, integrated model, integrated database)
Opportunities: Reduction in cost of decision-making, reduction in overall schedule, possible quality improvement, stronger team-building and “ownership” of engineering and design decisions and activities	Threats: Risk of loss of control by overall project management due to delegation of decision-making away from central office

2.5 Summary of sensitivity analysis results

The SWOT analysis of the AS IS model reaction to the impacts described above leads to the following conclusions: There are possibilities to increase efficiency of business (work) processes in project management and project execution. To realize the potential impact, BPR is needed for development and implementation of an effective TO BE model.

3 Re-engineering of business-processes and TO BE model creation

The goal of business-process re-engineering (BPR) is to fundamentally change the way the organization operates. Information technology shouldn’t be “changed for change’s sake” or rationale for deploying different software tools. It has to be revolution through evolution and improve the company global business in a variety of ways, including the following:

- Customer responsiveness - We have to stay as close to customer as possible. Many globalization obstacles have to be overcome to do this.
- Faster time to market - The race belongs to the swift. By eliminating needless or redundant processes, using extranet tools for engineering, a company can streamline the operations to bring products and services to market faster than before. Therefore, in general, projects must be completed as fast as possible.
- Building barriers to competitive entry - Integrated models and databases provide innovative means to stay close to the global customers and create effective deterrents to competition.
- Enhanced innovation - Improvements in operations offer the potential to enable organizations to innovate in new ways quickly and continuously.

There are numerous approaches to BPR. One of the potential techniques is presented below and it includes six stages as follows:

- 1) Build Awareness - A situational assessment and problem statement have to be done, related to operational performance, customers, employees, competitors, industry and market trends.
- 2) Create an AS IS model - Find and observe the flawed, sub-optimal, non-value-added processes that are stifling desired results. “What needs to be changed” needs to be defined.
- 3) Envision alternate approaches - Once the problem has been diagnosed, a BPR team should articulate a new strategic direction for processes and associated measurements as well as evaluate new business alternatives. Costs of BPR have to be estimated. Benefits like organization affected (customers, development, services), type of benefit (cost reduction, increased revenue), timing of benefit (immediate, first year, future) have to be estimated too.
- 4) Redesign the business process - An overview of solutions must be made as well as a vision statement and estimation of goals. A TO BE model has to be created. Solution details have to be defined like changes to the organization (people, culture, and training), changes to the work processes, changes to support systems, changes for computer applications, etc. Also, the solution

alternatives (even the do nothing scenario) have to be analyzed. Critical assumptions and risk assessments have to be done as well as a SWOT analysis or some other analysis technique.

- 5) Implement the TO BE model - This stage involves not only the new processes in a new management and Information Technology (IT) infrastructure, but also training and adoption as well. Both IT and management change implementation should be done in parallel, following a joint plan.
- 6) Re-evaluate new potential - Tuning new processes, and continuously building the organizational learning into business processes is a key stage of the BPR life cycle.

One of the key elements of BPR is the development and implementation of information systems. Some of the main ideas used in corporate information systems are developed using the process approach to business management. The authors have presented such ideas in a previous paper (see reference [2]). Reduction in project engineering costs will come as a result of BPR and the implementation of a Corporate Information System (CIS) containing an integrated database. This relationship of costs (and costs reduction between AS IS and TO BE models) is depicted in Figure 2 below. The curve represents the expected monthly cost of engineering. Note that for illustration purposes there are no units on "Cost" and "Time" but the idea is that implementation of BPR as mentioned in this paper will reduce the engineering cost of the project not only during the normal detailed "Engineering" stage, but also in at period which follows when "Correction" is made to account for mistakes, oversights and poor engineering quality. This "Correction" occurs during the construction period and only shows the engineering costs. The impact upon construction and procurement costs are much greater, but not within the scope of this discussion. To further clarify, the "Correction" shown in the chart is referring to activities such as correcting during the construction stage for inadequate engineering and design. It can be viewed as a time period when the project is under construction and either the engineering was adequate and the engineering cost was as expected as in the TO BE model, or the engineering and design was found to be lacking and the engineering cost increased during the construction ("Correction") stage.

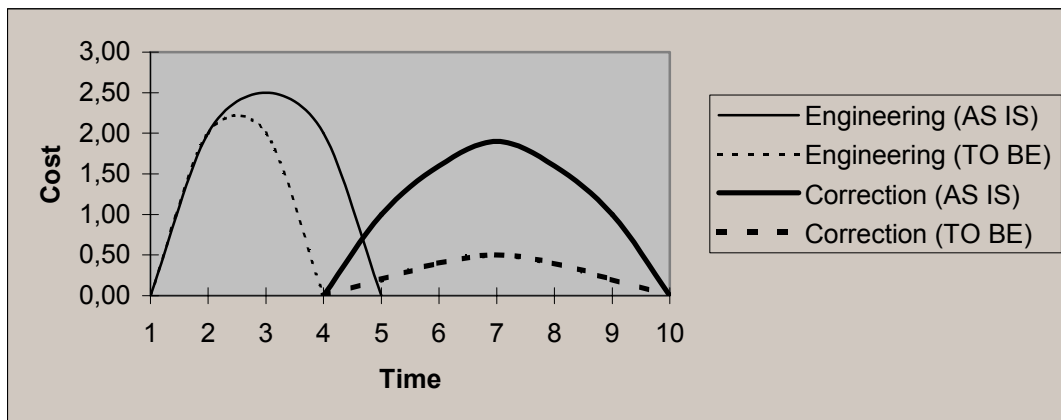


Figure 2 – Cost-Time curve after BPR

In general, the reason that there are significant cost reductions in the TO BE model is that there is higher quality project engineering and design due to implementation of concurrent engineering which utilized an effective extranet usage, an integrated design of facilities, a flattened organization structure, decentralization of decision-making, and a CIS for more effective management.

Conclusion

- 1) Problems that arise during international project development are highlighted.
- 2) An AS IS model is created and “fixed” for a sensitivity analysis.
- 3) Four impacts (changes) are applied separately to the AS IS model and reactions to these impacts are described.
- 4) A SWOT analysis is presented for each Impact (change).
- 5) A BPR technique is proposed to achieve strong points revealed in SWOT analyses and to develop a TO BE model.

Authors plan to develop results into methods for international project management.

References

- [1] V.A. Zarubin, H.R. Rigby, Globalization Problems in International Engineering Project Management, Engineering and Management Journal, Kazakhstan, 2002.
- [2] V.A. Zarubin, R.H. Rigby, Fundamentals of Process Approach to Open System Management // Proceedings of the V International Conference “Complex system: control and modeling problems”, June 17-21, 2003, Samara, Russia.
- [3] V.A. Vittikh, Engineering theories as a basis for integrating deep engineering knowledge // Artificial Intelligence in Engineering, v.11, n.1, 1997.
- [4] V.A. Vittikh, Concept of open organizational system management, Proceedings of Samara Center of RAS, 1999.

СИСТЕМНЫЙ АНАЛИЗ И МОДЕЛИ ИНФОРМАЦИОННОГО ВЗАИМОДЕЙСТВИЯ МЕХАНИЗМОВ ЭВОЛЮЦИОННОГО РАЗВИТИЯ

Н.В. Дилигенский¹, А.П. Ефимов²

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
iccs@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²Самарский государственный технический университет
443100, Самара, ул. Молодогвардейская, 244, Главный корпус СамГТУ, Россия
entcom@samgtu.ru

тел: +7 (8462) 32-42-34, факс: +7 (8462) 32-11-06

Ключевые слова: *Сложные открытые системы, эволюция, целенаправление, саморазвитие, саморегулирование, самоорганизация, информационное взаимодействие*

Abstract

Self-development, self-regulation and purposefulness models mechanisms, which information co-operation is defined by complex open systems evolution strategy are considered.

Введение

Процессы эволюции в сложных, открытых системах обуславливают их развитие, повышение разнообразия поведения, образование новых структур, улучшение качества функционирования. Они обеспечивают адаптацию существования систем во внешней, изменяющейся среде. Процессы эволюции являются результатом информационного взаимодействия трех системных механизмов развития: целенаправления, саморегулирования и саморазвития.

Целенаправление реализуется внешним воздействием, обеспечивающим изменение состояния системы в желаемом направлении развития. Механизм саморегулирования обеспечивает устойчивость, стабильность функционирования системы на основе отрицательной обратной связи, вырабатывающей внутренние управляющие воздействия, уменьшающие отклонения от желаемого направления развития и компенсирующие влияние всевозможных помех и возмущений. Саморазвитие обуславливается механизмом положительной обратной связи, вырабатывающим цели, определяемые внутренними, структурными и функциональными характеристиками системы. При этом внутренние цели системы, как правило, иные, чем внешние цели, формируемые механизмом целенаправленного развития.

При независимом, автономном действии каждый механизм управления играет принципиально различную роль в функционировании систем. Целенаправление определяет генеральную, стратегическую направленность развития системы. Механизм саморегулирования стабилизирует поведение, обеспечивает адаптацию к изменениям, уменьшает многообразие состояний систем, приводит к разрушению, дезорганизации структур. Механизм саморазвития приводит к росту многообразия, к образованию и организации новых структур и существованию неустойчивости, как общесистемного свойства.

1 Базовая модель механизмов эволюционного развития

В качестве прототипа развивающейся, эволюционной системы, характеризующейся взаимодействием трех вышеперечисленных механизмов управления, рассмотрим экономическую макросистему с агрегированным капитальным ресурсом. Базовую модель динамики экономи-

ческой системы запишем на основе баланса финансовых потоков в форме модифицированного уравнения Солоу [1].

$$(1) \quad Y + I = P + N,$$

где Y — продукция, произведенная системой, I — внешнее финансовое воздействие (внешние инвестиции), P — потребленная продукция, N — финансовые средства, идущие на развитие (внутренние инвестиции).

Произведенная продукция Y при стабильном уровне агрегированной технологии производственной деятельности определяется двухфакторной производственной функцией типа Кобба-Дугласа [2]

$$(2) \quad Y(t) = A \cdot K^\alpha(t) \cdot L^\beta(t),$$

где $K(t)$ — используемые капитальные ресурсы, $L(t)$ — трудовые ресурсы, задействованные в производственной деятельности, A , α , β — параметры модели (постоянные), характеризующие эффективность агрегированной технологии, t — время.

Потребление P в (1) полагается пропорциональным произведенной продукции

$$(3) \quad P = pY,$$

где p — коэффициент потребления.

Считается, что внутренние инвестиции N идут на возобновление израсходованных капитальных ресурсов и на увеличение капитала [1]

$$(4) \quad \frac{dK}{dt} + aK = N,$$

где a — коэффициент выбытия (расходования) капитала.

Совокупность соотношений (1)–(4) сводится к одному дифференциальному уравнению относительно капитальных ресурсов $K(t)$

$$(5) \quad K' = (1 - p)AL^\beta K^\alpha - aK + I.$$

Уравнение (5) является некоторой модификацией модели Солоу [1], и описывает эволюцию экономической системы при взаимодействии внешнего воздействия и внутренних механизмов самоорганизации. Первый член правой части уравнения (5) характеризует механизм саморазвития, реализуемый положительной обратной связью. Второй член отвечает механизму саморегулирования, обеспечивающему устойчивость функционирования системы на основе отрицательной обратной связи. Третий член описывает внешнее управляющее воздействие.

Положим в (5), что численность трудовых ресурсов $L(t)$ изменяется незначительно за анализируемый период времени T , и ее можно приближенно положить постоянной, равной среднему значению

$$(6) \quad \hat{L} = \frac{1}{T} \int_0^T L(t) dt.$$

В этом случае модель эволюции экономической системы (5) запишем в следующем виде относительно капитального ресурса системы K

$$(7) \quad K' = bK^\alpha - aK + I,$$

где $b = (1 - p)A\hat{L}^\beta$.

Дадим следующую расширенную интерпретацию полученного уравнения (7)., Учитывая симметричность (2) относительно K и L , рассуждая аналогичным образом, получим, что поведение другого — трудового ресурса $L(t)$ — описывается таким же как (7) уравнением. В соответствии с этим будем считать, что зависимость (2) можно также использовать для описания

функционирования развивающейся, эволюционной системы при воздействии иных факторов (ресурсов): природно-сырьевых, топливно-энергетических, информационных, инновационных, и других. Положим далее, что совокупность ресурсов эволюционной системы можно интегрировано охарактеризовать одним агрегированным обобщенным ресурсом X . Его величина получается путем свертывания определенным образом значений основных, базовых факторов. При такой интерпретации будем считать, что уравнение (7) пригодно для описания динамики обобщенного ресурса X эволюционной системы.

В соответствии с приведенными рассуждениями запишем уравнение (7) в следующем виде

$$(8) \quad X' = bX^\alpha - aX + I.$$

Далее уравнение (8) будем считать базовой агрегированной моделью, усреднено описывающей процессы эволюции в открытых, развивающихся системах. Характеристики и параметры базовой модели будем интерпретировать следующим образом.

Величину Y будем считать обобщенной координатой, интегрально характеризующей состояние (потенциал) системы. При этом зависимость Y от обобщенно ресурса X отвечает постоянной относительной (логарифмической) чувствительности состояния системы к изменению ресурсных воздействий — $\alpha = X dY/Y dX = const$. Из этого следует, что обобщенные координата и ресурс системы связаны элементарным соотношением $Y = BX^\alpha$, где B – масштабный показатель. Очевидно, легко записать уравнение эволюции для координаты состояния Y , однако оно более громоздкое и менее удобное для анализа, и в качестве базовой будем использовать модель (8).

Первый член в правой части (8) описывает механизм формирования внутренних целей системы. Параметр b является обобщенным показателем, интегрально характеризующим интенсивность образования внутренних целевых установок. Второй член (8) обеспечивает стабилизацию процессов. Обобщенный параметр a характеризует интегральные издержки на добычу, приобретение, преобразование и использование всех видов ресурсов системы и объединение (агрегирование) их в обобщенный ресурс X . I – внешнее целенаправляющее воздействие. Будем считать, что агрегированные характеристики системы $X(t)$ и $Y(t)$ являются наблюдаемыми и могут быть каким-либо способом измерены. По фактическим значениям $X(t)$ и $Y(t)$ далее могут быть идентифицированы параметры модели α, a, b .

Уравнение (8) записано в непрерывном времени. Данные наблюдений, как правило, существуют для некоторой последовательности моментов времени (лет или других единиц). С учетом этого обстоятельства запишем уравнение (8) для дискретного времени $t = n$

$$(9) \quad X_{n+1} = bX_n^\alpha - (a-1)X_n + I = f(a, b, I, \alpha, X_n), \quad n=0, 1, 2, \dots$$

где $f(\bullet)$ - функция последования, характеризующая закон дискретного изменения последовательных состояний эволюционной системы. Далее будем считать, что уравнениями (8), (9) можно описать процессы эволюции неживой и живой природы и общества в технических, эргатических, социально-экономических и экологических развивающихся системах.

2 Процессы эволюции согласованного развития

Исследуем возможные стратегии эволюционного развития систем, описываемых уравнениями (8), (9), в зависимости от интенсивности взаимодействующих механизмов саморегулирования, саморазвития и целенаправленного воздействия. На первой стадии изучим сценарии равновесного эволюционирования динамических систем и условия их существования. Равновесные состояния, которые обозначим через X^* , Y^* , отвечают стационарным решениям (8), (9) при t (или n) $\rightarrow \infty$.

Следовательно, X^* являются корнями алгебраического уравнения

$$(10) \quad aX^{*\alpha} - bX^* - I = 0.$$

Решения (10) определяют равновесные состояния эволюционных систем как функции характеристик механизмов саморегулирования a , саморазвития α и b и целенаправленности I . При известных значениях параметров a , α , b , I вычисление величины X^* по (10) может быть проведено различными численными методами. Условия наличия конструктивно интерпретируемых решений (10) – вещественных, положительных X^* – определяют условия существования равновесных состояний эволюционных систем.

По содержательному смыслу величина логарифмической функции чувствительности α является положительной величиной, меньшей единицы. Идентификация значений α в многочисленных конкретных ситуациях показывает, что её значения, как правило, лежат в диапазоне $0,3 \div 0,8$ [4], [5]. Далее, для получения качественно обозримых решений положим $\alpha=0,5$. Отметим, что при этом структурные свойства процессов эволюции неизменны в диапазоне изменений значений $\alpha \in (0, I)$.

Получая в этом случае аналитическое решение (10), найдем, что равновесные состояния системы существуют в области значений параметров:

$$(11) \quad I > -\frac{b^2}{4a}.$$

При этом величина обобщенного ресурса X^* равна:

$$(12) \quad X^* = \frac{1}{2a} \cdot \left(b^2 + 2aI + b\sqrt{b^2 + 4aI} \right).$$

Равновесное состояние системы Y^* характеризуется обобщенной фазовой переменной:

$$(13) \quad Y^* = \frac{b}{2a(1-p)} \cdot \left(b + \sqrt{b^2 + 4aI} \right).$$

Соотношения (12), (13) определяют односточечные аттракторы, притягивающие (или отталкивающие) временные траектории динамических систем.

Проанализируем характер эволюционирования систем в окрестности аттрактора (12). В соответствии с (9) динамика процессов однозначно определяется свойствами функции последования $f(a, \alpha, b, I, X)$. В зависимости от поведения f выделим два класса процессов с различными структурными свойствами.

При $\partial f / \partial X^* > 0$ функция последования определяет тенденцию прогрессивной эволюции, т.е. постоянного роста базовых показателей функционирования системы X^* и Y^* во времени. В этой области все три различных механизма эволюции – саморегулирования, саморазвития и целенаправленности – согласованы, сбалансированы между собой, и в системе протекают гладкие, монотонные процессы достижения равновесного состояния.

На рисунке 1 приведены временная траектория и фазовый портрет процессов эволюции для области согласованных механизмов управления. Поведение интегральных траекторий при этом отвечает фазовому портрету «устойчивый узел». Такие процессы согласованного эволюционного развития существует лишь в некоторой области притяжения односточечного аттрактора (12). Эта область является, по терминологии Н.Н. Моисеева «каналом эволюции» [6, 7], определяющим существование структурно-устойчивых стратегий эволюционирования развивающихся систем.

Границей канала согласованного эволюционного развития является условие $\partial f / \partial X^* = 0$. Проводя соответствующие вычисления в (9), для границ этого канала эволюции в пространстве параметров a, b, I получим описание:

$$(14) \quad I = \begin{cases} +\infty, & 0 < a < 1 \\ \frac{2-a}{(1-a)^2} \cdot \frac{b^2}{4}, & 1 < a < \infty \end{cases} .$$

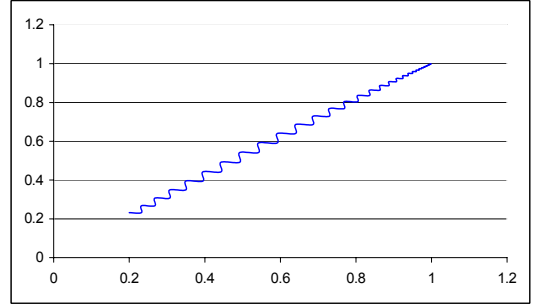
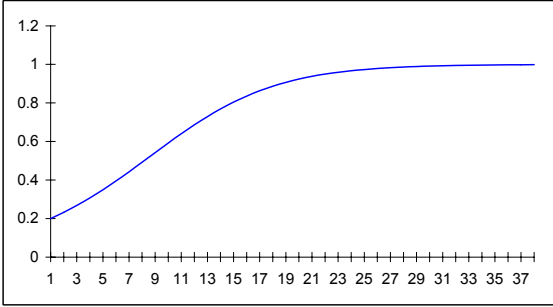


Рисунок 1

Другой границей канала является поверхность (11). На рисунке 2 представлены сечения эволюционного канала согласованного развития в плоскостях управляющих параметров a, b и a, I .

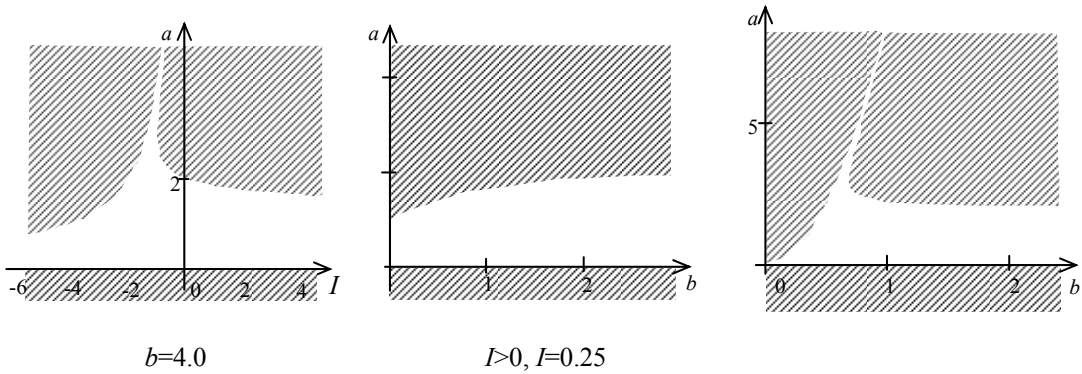


Рисунок 2 – Канал согласованного развития.

Видно, что согласование различных механизмов эволюции наиболее эффективно осуществляется в зоне малых интенсивностей саморегулирования. При этом с снижением a размеры области сбалансированного развития постоянно увеличиваются, становясь бесконечным при $a \rightarrow 0$. интенсивность целеполагания I и саморазвития b немонотонно влияют на размеры этой зоны, и существуют некоторые экстремальные значения I и b , при которых размеры области согласованного развития максимальны.

3 Равновесное, адаптивное, эволюционное развития

При $\partial f / \partial X^* < 0$ функция последования определяет тенденции негативного эволюционирования динамической системы. В этом случае увеличение базового ресурса, используемого

системой X_n , приводит к уменьшению последующего поступающего в распоряжение системы ресурса X_{n+1} и, следовательно, к снижению потенциала системы Y_{n+1} . Механизм самоорганизации работает в направлении, противоположном целенаправляемому развитию, и внутренние цели, формируемые контуром саморазвития, противоречат внешним целевым воздействиям. Сценарии эволюции в этом случае отвечают адаптационным процессам отыскания компромиссов между противоположными тенденциями развития.

Изучим условия сходимости последовательности компромиссов. В зависимости от глубины противоречий компромиссы могут реализоваться в различных формах. Известно [3], что процессы, описываемые дискретной рекуррентной последовательностью (9) сходятся, если выполняется условие:

$$(15) \quad -1 < \left. \frac{\partial f(a, \alpha, I, X)}{\partial X} \right|_{X=X^*} < 1.$$

Неравенство (15) определяет область существования устойчивых равновесных процессов в эволюционных системах, сходящихся к одноточечному аттрактору (12). Класс стратегий устойчивого, согласованного, эволюционного развития, отвечающих (15) при $\partial f / \partial X^* > 0$, был исследован выше, и его условия существования определяются соотношениями (11), (14) а также $a = 0$, отвечающими правой части неравенства (15). Границы области протекания сходящихся адаптационных процессов при $\partial f / \partial X^* < 0$, в соответствии с (15) определяются левой частью неравенства (15) $\partial f / \partial X^* = -1$ и поверхностью $\partial f / \partial X^* = 0$ (14). Проводя соответствующие вычисления в (9) при $\partial f / \partial X^* = -1$, получим следующее выражение для левой границы этой области с $\partial f / \partial X^* < 0$.

$$(16) \quad I = \begin{cases} \infty, & 0 < a < 2, \\ \frac{4-a}{(2-a)^2} \cdot \frac{b^2}{4}, & 2 < a < \infty. \end{cases}$$

Таким образом, устойчивый компромисс между противоположными целями самоорганизации и целенаправляемого развития реализуется в области, заключенными между границами (14) и (16).

Конфигурации зон существования этого эволюционного канала приведены на рисунке 3.

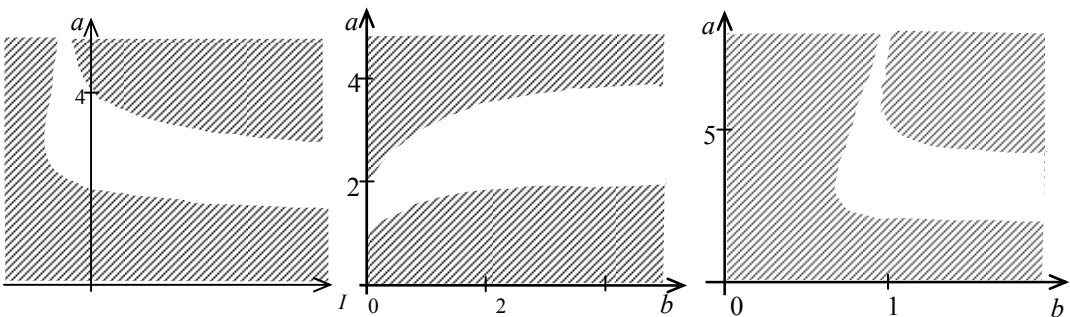


Рисунок 3 – Канал равновесного адаптационного развития.

Отчетливо видна структуризация обеих границ канала согласованного эволюционного развития и резкое сужение его русла при увеличении a .

На рисунке 4 представлены временная траектория и фазовый портрет процессов эволюции для этого случая. Видно, что в этом канале эволюции процессы самоорганизации реализуются в форме колебательных, сходящихся режимов. Поведение интегральных траекторий отвечает фазовому портрету «устойчивый фокус», и стационарным равновесным состоянием является одноточечный, устойчивый аттрактор.

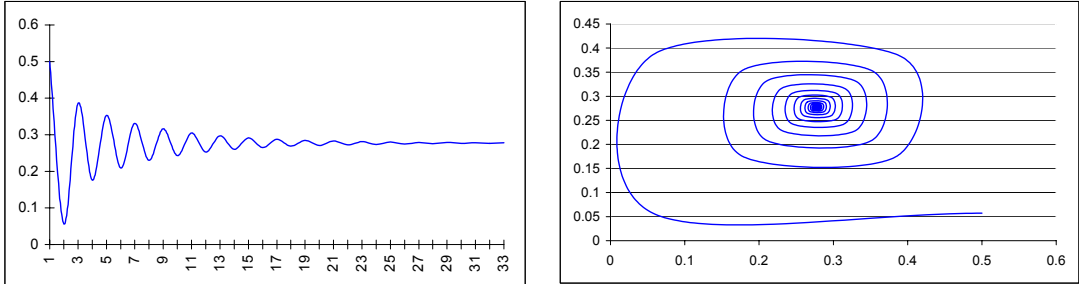


Рисунок 4

4 Процессы эволюции циклического развития

При бóльших рассогласованиях внутренних и внешних целей системы при $\partial f / \partial X^* < -1$, противоречия между различными механизмами эволюционного развития настолько возрастают, что оказываются неразрешимыми в классах равновесных состояний. Стационарные решения уравнения (9), отвечающие одноточечному аттрактору, оказываются неустойчивыми. При $\partial f / \partial X^* = -1$ решения разветвляются [3, 8] происходит бифуркация, и из одноточечного аттрактора рождается двухточечный. Система переходит в другой канал эволюционного развития в окрестности образовавшегося двухточечного аттрактора.

Характеристики двухточечного аттрактора находятся как неподвижные точки двукратно итерированного дискретного отображения (9), т.е. как решения уравнения

$$(17) \quad X^* = f(f(a, b, I, X^*)) \equiv f^2(a, b, I, X^*).$$

Двухточечный аттрактор является устойчивым в области, левой границей которой является поверхность потери устойчивости одноточечного аттрактора (16), а правая граница, аналогично (15), отыскивается из условия $\partial f^2 / \partial X^* = -1$. В этом канале эволюции двухточечный аттрактор формирует простой, устойчивый, предельный цикл [8, 9]. Соответствующие ему стратегии эволюционного развития отвечают сходящимся процессам динамической самоорганизации — устойчивым автоколебательным режимам.

Процессы эволюции имеют характер циклического развития — фаза роста и подъема характеристик системы сменяется этапом спада, уменьшения показателей, и так далее. Эволюционные процессы и фазовые портреты состояний систем в этом канале циклического эволюционного развития приведены на рисунке 5. Конфигурации границ канала аналогичны рассмотренной выше области равновесного, адаптационного развития, при этом диапазон характеристик механизмов эволюции, в которой существует этот класс эволюционных стратегий, уменьшается.

При переходе через найденную правую границу устойчивости этого канала эволюции, отвечающему увеличению a и I и уменьшению b , происходит следующая бифуркация, и из двухточечного аттрактора формируется четырехточечный. Его характеристики находятся как решения уравнения

$$(18) X^* = f(f(f(f(a,b,I,X^*))) \equiv f^4(a,b,I,X^*).$$

Соответствующие аттрактору (18) эволюционные процессы отвечают двойному предельному циклу. На рисунке 5 приведены переходный процесс и фазовый портрет, отвечающие такому каналу циклического эволюционного развития.

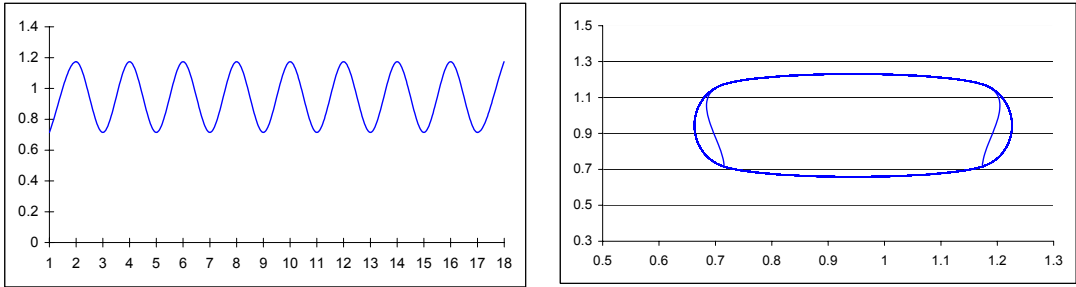


Рисунок 5

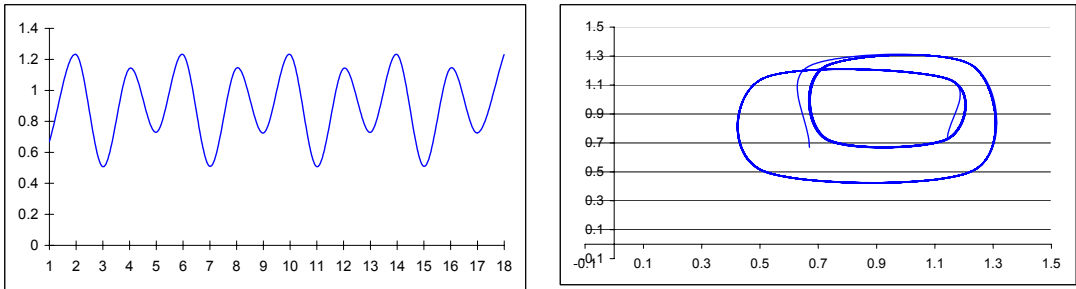


Рисунок 6

При дальнейшем возрастании a и I и уменьшении b происходит следующая бифуркация, образуется восьмиточечный аттрактор, затем новая бифуркация и шестнадцатиточечный аттрактор и так далее. Эволюционные процессы при этом отвечают большему многообразию самоорганизующихся режимов циклического развития с высоким числом степеней свободы и с сложными предельными циклами.

5 Канал хаотической эволюции

При значительной кратности предельных циклов соседние ветви периодических аттракторов находятся весьма близко друг к другу, и при малейших флуктуациях изображающие точки состояний систем случайным образом переходят с одной ветви на другую. В системе начинают протекать неустойчивые, с высокой степени неопределенности процессы самоорганизации эволюционирования. На рисунке 7 приведены временная траектория и фазовый портрет, отвечающие такой ситуации. Видно отсутствие закономерности и прогнозируемости протекания процессов. Такой канал эволюционного развития отвечает странным аттракторам с конечным числом степеней свободы типа Лоренца [10]. В этом случае противоречия взаимодействующих механизмов эволюции — самоорганизации и целенаправления — оказываются настолько большими, что неразрешимы в рамках регулярного поведения систем.

Затем число ветвей в каскаде бифуркаций становится бесконечно большим, процессы становятся полностью хаотическими, отвечающие стохастическим, фрактальным, бесконечномер-

ным аттрактором (рисунок 8). Эволюционирование систем оказывается непредсказуемым. В системе наступает хаос, при этом он является структурированным, с детерминированными зонами самоорганизации [8-12].

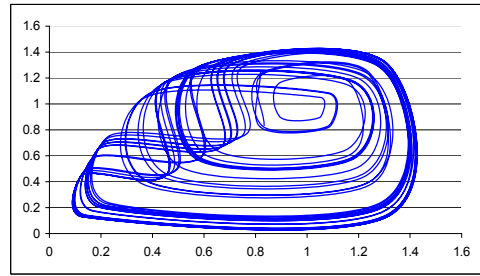
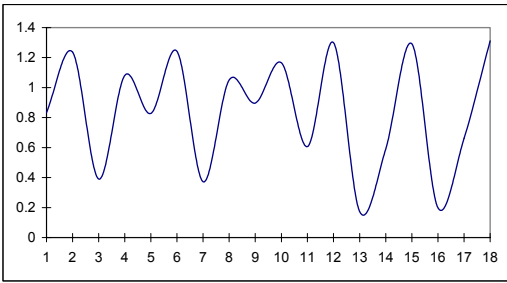


Рисунок 7

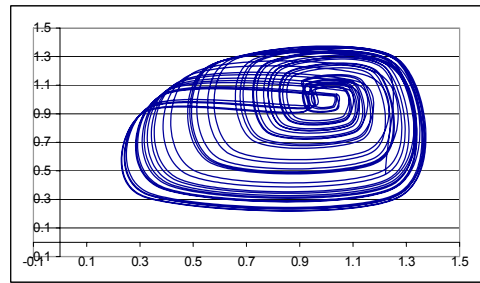
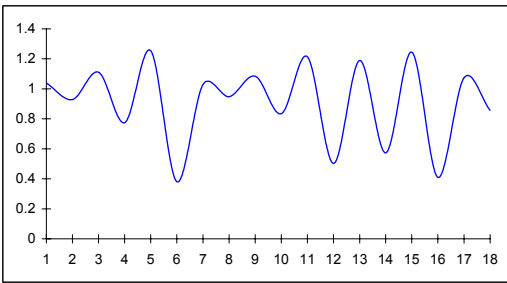


Рисунок 8

6 Трансформация каналов эволюционного развития

Процессы перехода систем из одного канала эволюционного развития в другой происходят по разным сценариям. Для анализа влияния каждого из механизмов эволюции на характер переходов построим расширенные фазовые портреты, где одной из осей будут особые, предельные при $n \rightarrow \infty$ значения координаты состояния Y , а другой — каждый из обобщенных параметров a, b, I при фиксированных величинах двух других.

Общесистемной закономерностью процессов эволюции является повышение системной эффективности использования ресурсов (возрастание b) и увеличение затрат на приобретение ресурсов – рост a .

На рисунке 9 приведен расширенный фазовый портрет в координатах Y, a . Видно, что он имеет вид бифуркационной диаграммы, графически иллюстрирующей ветвления особых состояний систем.

На рисунке 9 приведены также полученные путем реализации численных экспериментов инверсные значения показателя Ляпунова L , которые приняты в качестве характеристики устойчивости временных траекторий [9]:

$$(19) L(a, b, I) = \lim_{m \rightarrow \infty} \frac{1}{M} \cdot \sum_{m=0}^M \ln \left| \frac{\partial f(a, b, I, X)}{\partial X} \right|_{X=X_m},$$

где X_m результат реализации численного эксперимента для m -го последовательного отображения зависимости (9) $X_m = f^m(a, b, I, X_0)$, X_0 - начальное состояние, M - число экспериментов. Из бифуркационной диаграммы видна последовательная смена при возрастании a разных кана-

лов эволюционного развития: согласованного; равновесного, адаптивного: циклического; хаотического. В целом, этот сценарий эволюционного развития отвечает теории универсального перехода к хаосу Фейгенбаума [11, 12].

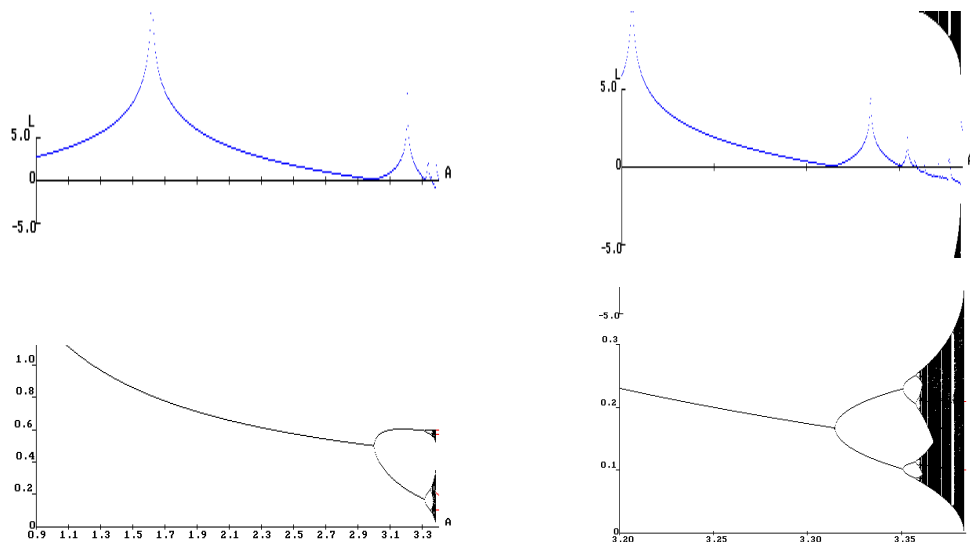


Рисунок 9

На рисунке 10 приведен сценарий эволюционного развития в координатах Y, b . Видно, что он также характеризуется бифуркационной диаграммой, однако в этом случае разветвление состояний, организация каскада бифуркаций и хаотизация эволюционирования происходят при уменьшении b .

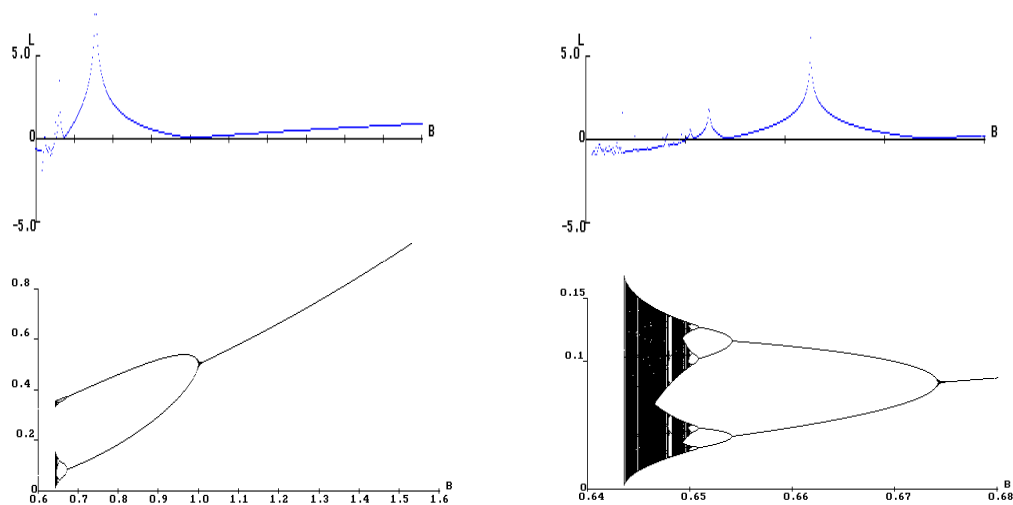


Рисунок 10

Сценарий эволюционного развития в зависимости от интенсивности внешнего, направляющего воздействия I приведен на рисунке 11.

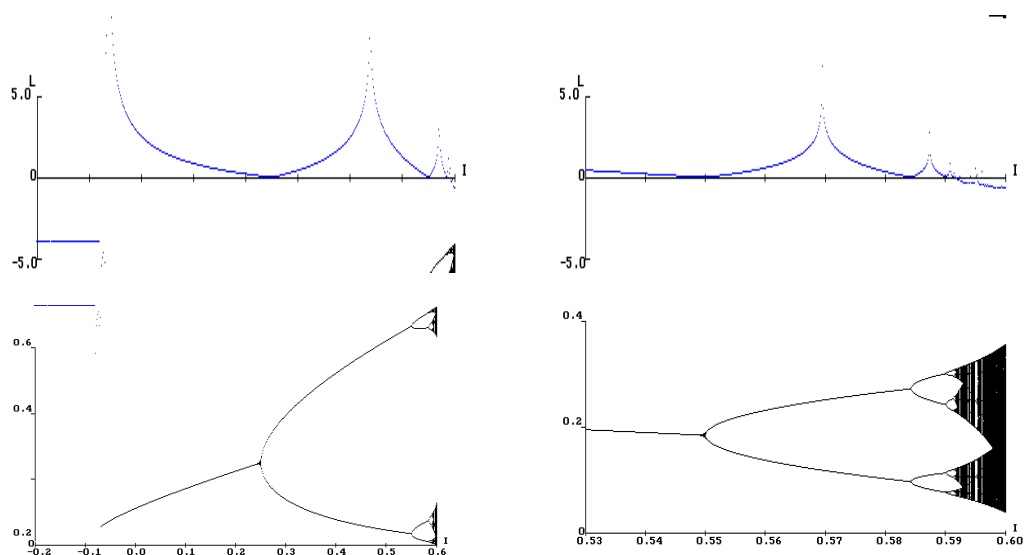


Рисунок 11

Видно, что он также имеет характер бифуркационной диаграммы. При уменьшении I в некоторой, вполне определенной точке происходит катастрофа (в смысле математической теории катастроф Тома [13]). В рассматриваемой ситуации катастрофой является складка Уитни [14], и в точке катастрофы система мгновенно теряет свою целостность, прекращает существование, рассыпается.

Заключение

Проведенный модельный анализ выявил многообразие сценариев эволюционного развития и геометрию каналов эволюции сложных развивающихся систем, в целом, соответствующих парадигме универсального эволюционизма Н.Н. Моисеева [6, 7], и определяющих общественные, глобальные закономерности стратегий обеспечения устойчивого развития (*sustainable development*) и поддержания «устойчивого неравновесия» (*sustainability*).

Список литературы

- [1] Аллен Р. Математическая экономия. М.: И.Л., 1963.
- [2] Клейнер Г.Б. Производственные функции: теория, методы применения. М.: Финансы и статистика, 1986. 239 с.
- [3] Баутин Н.Н., Леонтович Е.А. Методы и приемы качественного исследования динамических систем на плоскости. М.: Наука, 1976. 496 с.
- [4] Гранберг А.Г. Моделирование социалистической экономики. М.: Экономика. 1988. 307 с.
- [5] Дилигенский Н.В., Цапенко М.В. Построение и анализ макроэкономических моделей регионального производственно-технологического комплекса / Высшее образование, бизнес, предпринимательство // Материалы Межвузовского сборника научных трудов. Самара. СамГТУ, ПИБ. 2000. с.149-155.
- [6] Моисеев Н.Н. Человек и ноосфера. М.: Молодая гвардия. 1990. 352 с.
- [7] Моисеев Н.Н. Универсум. Информация. Общество. М.: Устойчивый мир. 2001. 200 с.

- [8] Дилигенский Н.В., Ефимов А.П. Сингулярное вырождение и хаотическое поведение однопараметрических эволюционных систем // Проблемы моделирования и управления в сложных системах. Самара. СНЦ РАН, 2002. С. 91-100.
- [9] Дилигенский Н.В., Ефимов А.П. Модельный анализ самоорганизации, порядка и хаоса в развивающихся системах // Проблемы моделирования и управления в сложных системах. Самара. СНЦ РАН, 2003. С. 115-122.
- [10] Lorenz E.N. Deterministic nonperiodic flow. J. Atmosferic. Sci. 1963.V.20.W2. p. 130-141.
- [11] Feigeubaum M.J. Quantitative universality for a class of nonlinear transformations. J.Stat.Phys. 1978.V.19W1. p 25-52.
- [12] Фейгенбаум М. Универсальное поведение в нелинейных системах. // Успехи физ. наук, 1983. Т141, вып.2. С. 343-374.
- [13] Thom R. Catastrophe Theory: Its present state and future perspectives. // Dinamical Systems. Warwick, 1974 –Berlin – Heidelberg –New York; Springer – Verlag, 1-75.- p. 366-372. Lecture Notes Math.V.468.
- [14] Арнольд В.Н. Теория катастроф. М.: Наука, 1990. 128 с.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ОБОБЩЁННОЕ ОЦЕНИВАНИЕ ЭФФЕКТИВНОСТИ ПРОИЗВОДСТВЕННО-ЭКОНОМИЧЕСКИХ СИСТЕМ

Н.В. Дилигенский¹, М.В. Цапенко²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61

iccs@iccs.ru

тел: +7 (8462) 33-26-77, факс: +7 (8462) 33-27-70

²Самарский государственный технический университет

443010, Самара, ул. Молодогвардейская, 244

usat@samgtu.ru

тел: +7 (8462) 32-42-34

Ключевые слова: системная эффективность функционирования, показатели эффективности, сравнительный анализ, производственные функции, методология Data Envelopment Analysis (DEA), метод анализа иерархий (Analytic Hierarchy Process)

Abstract

We analyzed models and methods socio-economic systems effectiveness generalized, multi-criterion estimations construction based on different means of aggregation local criterions of functioning efficiency, formal structures and expert conclusions.

Введение

Важнейшими задачами моделирования и управления сложными системами являются повышение системной эффективности их функционирования – оптимальное использование имеющихся в распоряжении системы ресурсов для достижения высоких конечных результатов. В этой связи актуальными являются разработка и применение математических моделей и системных методов оценивания эффективности, проведения сравнительного анализа, сопоставления исследуемых объектов, определения и выбора приоритетов. Объектами, для совершенствования которых возможно конструктивное применение этих методов, являются различные по классу, масштабу и составу производственные, социальные и экономические системы.

В статье в качестве базовой будем рассматривать многомерную производственно-экономическую систему, имеющую n – входных факторов (соответствующих ресурсов) и m – выходных величин, характеризующих выпуск продукции. Каждую из входных и выходных величин будем считать локальным частным критерием качества, вносящим соответствующий вклад в обобщённый показатель эффективности. В соответствии с этим, глобальный критерий качества является многокритериальным, представляющим композицию (объединение) $n+m$ частных критериев. При этом выходные характеристики связаны прямой, а входные обратной зависимостью с обобщённой системной целевой функцией.

Существуют различные способы построения моделей и получения системных оценок эффективности. Одни основаны на применении только формальных математических подходов – методов теории производственных функций, математического программирования (Data Envelopment Analysis). Другие используют, наряду с модельными конструкциями дополнительно экспертные оценки, например, метод анализа иерархий (МАИ).

Комплексное применение совокупности этих методов позволяет формировать обобщённые, глобальные критерии эффективности, проводить сопоставление сложных систем, определять приоритеты и направления будущего развития исследуемых объектов.

1 Модели оценки эффективности на основе аппарата производственных функций

В случаях если сложная система представляет собой процесс преобразования некоторого множества входных факторов – ресурсов в комплекс выходных характеристик – конечных продуктов, т.е. является производственной системой, то одним из наиболее распространённых видов математических моделей для описания функционирования и анализа эффективности этих систем является класс производственных функций (ПФ). Этот комплекс моделей представляет собой достаточный инструментарий для определения различных параметров рассматриваемой системы, таких как: необходимые объёмы производственных ресурсов, эффективность их использования, возможность взаимозамещения, определение дополнительно необходимых ресурсов и, в конечном счёте, прогнозирование выходных характеристик системы.

В общем виде ПФ представляет собой формализованное описание связи между множествами используемых ресурсов и выпускаемой продукции [1]:

$$(1) F(X, Y, A) = 0,$$

где:

X – вектор входных ресурсов;

Y – вектор выходных величин, выпускаемых продуктов;

A – вектор параметров ПФ;

F – оператор, описывающий базовые свойства и закономерности исследуемой экономической системы.

Выражение ПФ (1) можно записать в явной форме, разрешённой относительно Y :

$$(2) Y = f(X, A).$$

В (2) f – функция, определяющая зависимость выпуска продукции от входных ресурсов. Уравнение (2), описывающее протекание производственных процессов в сложной системе, является многомерной производственной функцией – МПФ [2]. Функциональная модель сложных систем, описываемых этим классом моделей, представлена на рисунке 1.

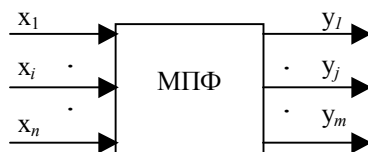


Рисунок 1 – Функциональная модель МПФ.

Применение моделей этого класса позволяет определить характеристики эффективности производственного процесса через соответствующие параметры ПФ и возможное прогнозирование объёмов выпуска продукции с учётом изменения входных характеристик системы.

Рассмотрим пример построения ПФ и анализа состава факторов системной эффективности на основе показателей эластичности. В качестве объекта исследования, математического моделирования и оценивания эффективности функционирования выберем региональный производственно-технологический комплекс Самарской области (РПТК). Моделируемый объект будем рассматривать как единое целое, без учёта его структуры, формы и способов организации производственных отношений.

Производственные ресурсы будем предельно агрегировать, и характеризовать двумя макроэкономическими показателями: количеством овеществлённого, прошлого труда (основными

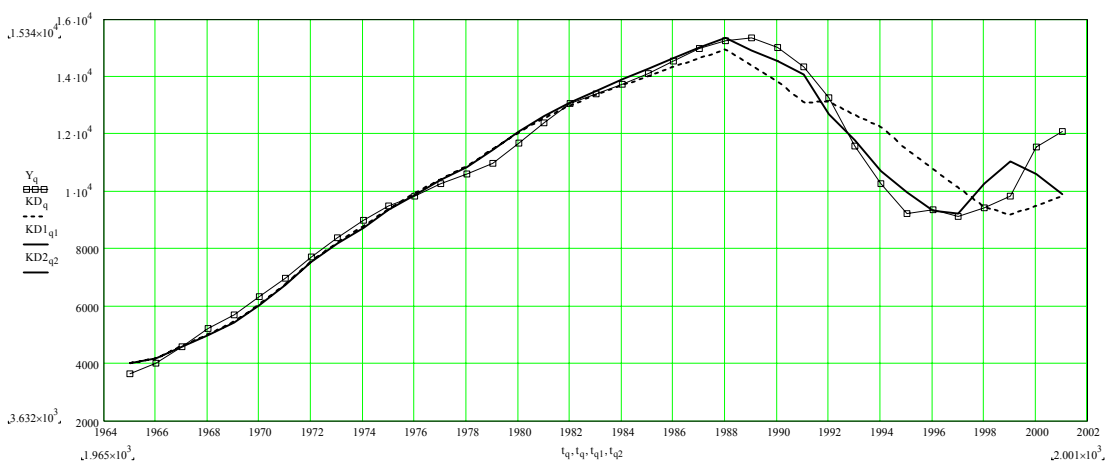
производственными фондами – K) и вкладом живого труда (численностью трудовых ресурсов – L). Исходные данные для построения возьмём из серии статистических ежегодников [3, 4] и работы [5]. Сопоставимость экономических характеристик в различных периодах обеспечим их пересчётом к единому уровню цен 1990 г. на основе фактических цен и приростных оценок показателей динамики производства.

В качестве базовой модели используем неоднородную производственную функцию Кобба-Дугласа:

$$(3) \quad Y(t) = A \cdot K(t)^\alpha \cdot L(t)^\beta, \quad \alpha + \beta \neq 1,$$

где A – коэффициент масштабной эффективности, α и β – факторные эластичности, характеризующие вклад соответствующих ресурсов в выпуск продукции.

Построим комплекс моделей на основе ПФ вида (3), при моделировании будем учитывать три периода развития экономической системы: общий – с 1965 по 2001 гг.; первый – с 1965 по 1990 гг.; второй – с 1990 по 2001 гг. Результаты моделирования представлены на рисунке 2.



- реальный выпуск продукции Y (тыс. руб. в сопоставимых ценах);
- расчёт по модели () на всём интервале исследования;
- расчёты по модели () для первого и второго интервалов).

Рисунок 2 – Траектории выпуска продукции в экономической системе (сглаженные данные).

На рисунке 2 наглядно видна сходимость модельных и реальных траекторий функционирования РПТК Самарской области. Анализ качественных характеристик моделей показал, что аппроксимативные свойства модели на первом интервале достаточно хорошие: $R^2=0.9956$, а на общем интервале близки к удовлетворительным: $R^2=0.9304$, что свидетельствует о высокой степени адекватности полученных моделей. На втором интервале модель практически непригодна для описания исследуемых процессов: $R^2=0.7689$.

Значениями эластичности выпуска по капиталу α , являются – на общем интервале – 0.59, на первом – 0.66. Эти значения, в целом, соответствуют современным данным об эффективном протекании процессов в экономической системе [2, 6]. Величина α на втором интервале имеет значение минус 3.19, что свидетельствует о деструктивном характере протекающих процессов в экономической системе в данный период.

Значения эластичности выпуска по труду β на общем и первом интервалах составляют 1.39 и 1.17, соответственно. Эти данные свидетельствуют о преобладающей роли трудовых ресурсов в производственной деятельности, и о недостаточном вкладе капитального фактора.

Иными словами, характер производственного процесса является трудозатратным [7]. На втором интервале значения β также принимают отрицательные значения (минус 0.95), что характеризует неудовлетворительное использование трудовых ресурсов.

На основе идентифицированной модели ПФ (3) определяется обобщённая системная эффективность РПТК, в целом. Общая эластичность выпуска по затраченным входным ресурсам $\lambda = \alpha + \beta$ показывает интегральную эффективность использования обеих ресурсов. Значения их на общем и первом интервалах равны 1.98 и 1.83, что свидетельствует об эффективности использовании ресурсов с расширением масштаба производства [1]. Отрицательное значение λ на втором интервале – минус 4.14 говорит о процессах деградации в экономической системе.

Взвешивая вклады относительных изменений ресурсов $K^* = K/K_0$, $L^* = L/L_0$ с нормированными эластичностями $\rho = \alpha / (\alpha + \beta)$, $1 - \rho = \beta / (\alpha + \beta)$, найдём величину интегрального масштабного эффекта M от совместного использования вовлечённых дополнительных капитальных и трудовых ресурсов в производственный процесс [7]:

$$(4) \quad M = K^{*\rho} \cdot L^{*(1-\rho)}.$$

Аналогично взвешивая относительную производительность каждого из ресурсов, получим обобщённый показатель технологической эффективности производства [7]:

$$(5) \quad R = \left(\frac{Y^*}{K^*} \right)^\rho \cdot \left(\frac{Y^*}{L^*} \right)^{1-\rho},$$

где $Y^* = Y/Y_0$; K_0 , L_0 , Y_0 – начальные значения факторов.

На рисунке 3 приведены значения интегрального масштабного эффекта – M , обобщённого показателя технологической эффективности – R и параметра относительной эффективности R/M .

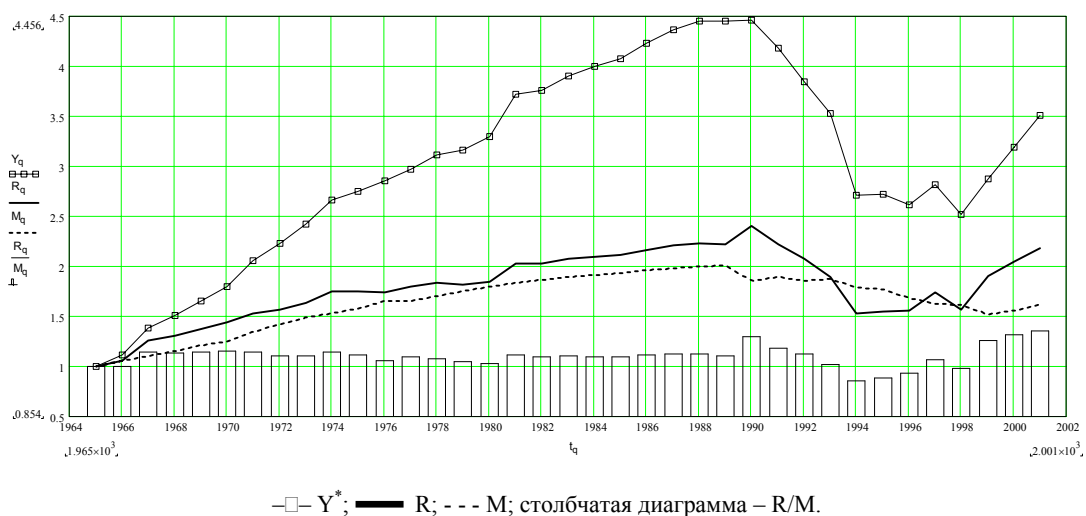


Рисунок 3 – Оценки параметров состава фактора эффективности.

Из рисунка 3 видно, что в период 1965 - 1990 гг. обобщённые показатели как масштабной, так и технологической эффективности РПТК постоянно росли. При этом технологическая эффективность R несколько превышала масштабную M . В 1990 году, за счёт совершенствования производственных технологий произошло увеличение выпуска продукции в 2,5 раза, а за счёт масштабного фактора в 1,8 раза.

На периоде с 1990 по 1998/99 гг. величины R и M падают, после 1999 года спад сменяется ростом. Так на конец интервала – 2001 г. из общего увеличения выпуска, изменения за счёт

технологической составляющей произошло в 2,2 раза, а за счёт расширения масштабов производства в 1,6 раза.

Относительная эффективность совершенствования технологических процессов в форме отношения технологического фактора к масштабному – R/M, несколько больше единицы (среднее значение 1,096), за исключением 1994, 1995, 1996 и 1998 гг., где доминирование масштабного роста объясняется большей степенью спада технологического фактора.

Таким образом, для РПТК Самарской области в период с 1965 по 2001 гг. изменения объёмов выпуска конечной продукции происходили, как за счёт расширения масштабов производства, так и посредством совершенствования технологических процессов, с некоторым преобладанием вклада последних.

Видно, что аппарат производственных функций позволяет построить достаточно адекватные математические модели реальных производственно-экономических систем и получать обобщённые, агрегированные оценки эффективности их функционирования.

2 Модели оценки сравнительной эффективности на основе методологии Data Envelopment Analysis

В предыдущем разделе были рассмотрены вопросы моделирования и оценивания обобщённой эффективности производственно-экономических систем на основе производственных функций.

Рассмотрим иной подход многокритериального оценивания сравнительной эффективности – методологию Data Envelopment Analysis (DEA) [8, 9]. Его существенным достоинством является минимальное присутствие субъективного фактора при формировании интегральных оценок эффективности.

Подход позволяет обоснованно, с учётом значительного числа входных и выходных параметров, в качестве которых могут выступать экономические, технологические, технические и другие критерии, находить обобщённые оценки эффективности процессов и производств в комплексе единой производственно-экономической, организационной системы.

Методы DEA являются обобщением расчётов использования входных ресурсных воздействий на многомерный случай. Для многомерного производственно-экономического объекта с m входными и k выходными параметрами: $X = \{X_1, X_2, \dots, X_m\}$ и $Y = \{Y_1, Y_2, \dots, Y_k\}$.

Комплексный обобщённый показатель эффективности для каждого n -го объекта в базовом варианте метода DEA формируется следующим образом:

$$(6) \quad f_n = \max_{u, v} \frac{u_{1n} \cdot Y_{1n} + u_{2n} \cdot Y_{2n} + u_{3n} \cdot Y_{3n} + \dots + u_{kn} \cdot Y_{kn}}{v_{1n} \cdot X_{1n} + v_{2n} \cdot X_{2n} + v_{3n} \cdot X_{3n} + \dots + v_{mn} \cdot X_{mn}}$$

В (6) u_{in} и v_{jn} – весовые коэффициенты, удовлетворяющие системе ограничений, определяющих область значений G :

$$(7) \quad \begin{aligned} \frac{u_{1n} \cdot Y_{11} + u_{2n} \cdot Y_{21} + u_{3n} \cdot Y_{31} + \dots + u_{kn} \cdot Y_{k1}}{v_{1n} \cdot X_{11} + v_{2n} \cdot X_{21} + v_{3n} \cdot X_{31} + \dots + v_{mn} \cdot X_{m1}} &\leq 1 \\ \frac{u_{1n} \cdot Y_{12} + u_{2n} \cdot Y_{22} + u_{3n} \cdot Y_{32} + \dots + u_{kn} \cdot Y_{k2}}{v_{1n} \cdot X_{12} + v_{2n} \cdot X_{22} + v_{3n} \cdot X_{32} + \dots + v_{mn} \cdot X_{m2}} &\leq 1 \\ &\dots \\ \frac{u_{1n} \cdot Y_{1N} + u_{2n} \cdot Y_{2N} + u_{3n} \cdot Y_{3N} + \dots + u_{kn} \cdot Y_{kN}}{v_{1n} \cdot X_{1N} + v_{2n} \cdot X_{2N} + v_{3n} \cdot X_{3N} + \dots + v_{mn} \cdot X_{mN}} &\leq 1, \end{aligned} \quad , \quad i = \{1, 2 \dots k\}; \quad j = \{1, 2 \dots m\}.$$

Система соотношений (6), (7) для $n=\{1, 2, \dots, N\}$ определяет N задач математического программирования. Решение каждой n -ой задачи ($n = 1, 2, \dots, N$) для n -ого объекта даёт значение соответствующего обобщённого показателя эффективности f_n , ранжированное на единичном интервале $[0, 1]$, и соответствующий ему набор весовых коэффициентов $u_n=\{u_{1n}, u_{2n} \dots u_{kn}\}$ и $v_n=\{v_{1n}, v_{2n} \dots v_{mn}\}$, максимизирующий функционал (6). Вообще говоря, эти веса u_{in} и v_{jn} будут различными для различных n -ых систем.

Отметим, что, в общем случае, решение задачи многокритериальной оптимизации (6), (7) имеет не единственное, а целое множество решений, удовлетворяющих заданным ограничениям (7) [10]. При этом, одной и той же относительной эффективности может отвечать некоторое множество решений как в пространстве входных и выходных параметров, так и в пространстве весов функционала. Следовательно, различные объекты с различающимися характеристиками в методе DEA могут иметь одинаковые обобщённые показатели эффективности.

Рассмотрим решение задачи определения сравнительной отраслевой эффективности производств, входящих в состав промышленного комплекса Самарской области. Оценку будем проводить на интервале, соответствующем переходному периоду развития: с 1990 по 2001 гг. Источником информации являются официальные материалы Самарского областного комитета государственной статистики [3].

В качестве модельного представления оцениваемых производств воспользуемся принятой ранее постановкой – рисунок 1. Каждую отрасль промышленности региона будем рассматривать в виде производственно-экономической системы, осуществляющей процесс преобразования основных входных ресурсов: капитальных (K) и трудовых (L), в агрегированный конечный продукт (Y).

В соответствии с методологией DEA, функционал оценки сравнительной эффективности представим в виде взвешенного отношения выходной характеристики – объёмного показателя выпуска продукции – Y к сумме взвешенных ресурсов: капитальных – K и трудовых – L. Таким образом, критерий оценки отраслевой сравнительной эффективности примет вид:

$$(8) \quad f_n = \max_{u, v \in G} \frac{u_n \cdot Y_n}{v_{1n} \cdot K_n + v_{2n} \cdot L_n},$$

где Y – объём продукции, произведённой в отрасли, млрд. руб. (для 1998 – 2001 гг. – млн. руб.); K – фонды основного вида деятельности (на конец года, по балансовой стоимости; млрд. рублей); L – численность персонала основной деятельности, тыс. человек; u, v_1, v_2 – положительные весовые коэффициенты.

Рассчитаем значения оценки сравнительной эффективности отдельно для каждого года по совокупности девяти региональных производств. В соответствии с методологией DEA для каждой отрасли в каждом году сформулируем задачу математического программирования вида (6), (7) и решим её. Таким образом, нахождение DEA-оценок отраслевой сравнительной эффективности в соответствии с функционалом (8) сводится к решению девяти оптимизационных задач на двенадцати интервалах исследования.

Результаты расчётов DEA-оценок сравнительной эффективности комплексов визуально представлены на графиках рисунка 4.

По результатам расчётов видно, что максимальную сравнительную эффективность на всём промежутке исследования имела отрасль пищевой промышленности. Этот отраслевой производственный комплекс занимал четвёртое место по валовому выпуску продукции во все годы (за исключением 2000 г.), но имел почти на порядок меньшие затраты основных фондов. В соответствии с этим, взвешенные интегральные затраты ресурсов у него наименьшие по отношению к другим комплексам, а его относительная эффективность функционирования максимальная. Все другие комплексы имели меньшую сравнительную эффективность в различные годы анализируемого периода.

На рисунке 5 приведены интегральные оценки эффективности функционирования промышленных комплексов за все годы.

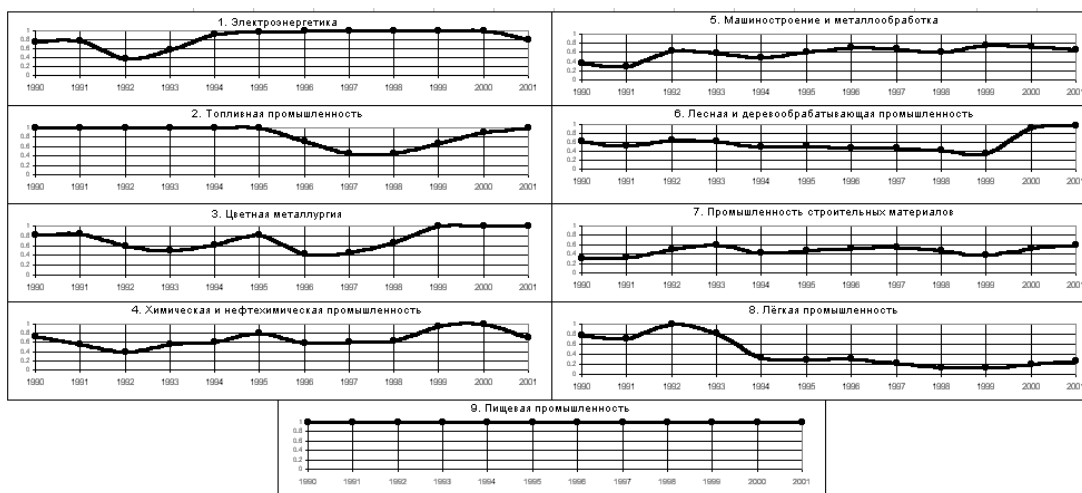


Рисунок 4 – Численные значения оценки отраслевой сравнительной эффективности

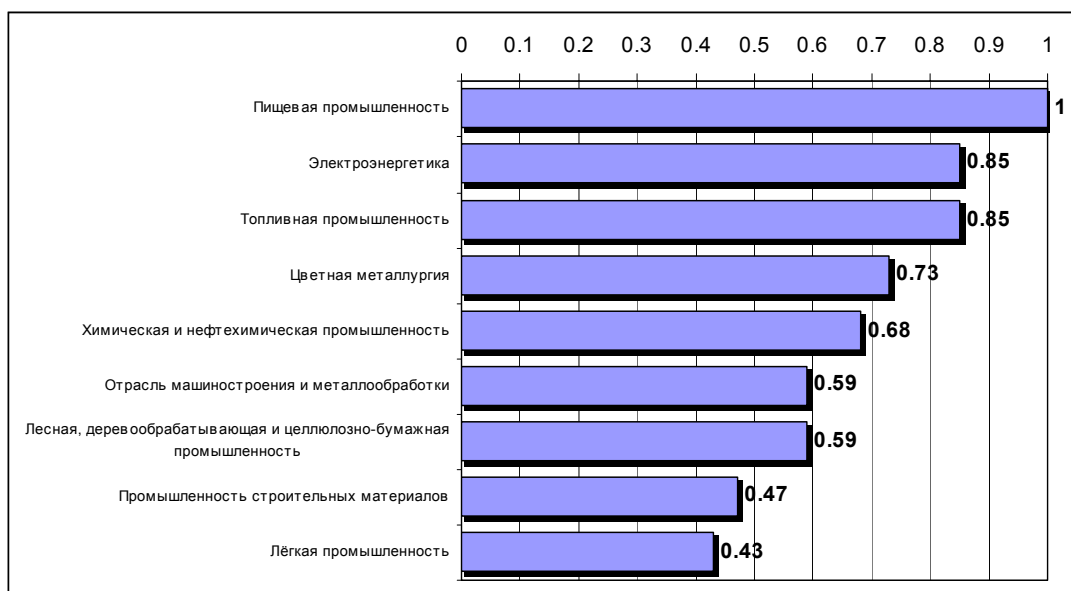


Рисунок 5 – Сводный рейтинг ведущих отраслей РПТК Самарской области.

Видна информативность полученных обобщённых оценок сравнительной эффективности использования комплексами имеющихся в их распоряжении ресурсов.

Таким образом, применённый метод DEA, реализующий формализованную процедуру свертывания известных локальных критериев эффективности в глобальный без использования дополнительной, как правило, субъективной информации о рангах частных показателей, является работоспособным для получения объективных заключений об обобщённой эффективности производств.

3 Экспертные модели оценки на основе метода анализа иерархий

В широком круге задач обобщённого оценивания являются неизвестными сами первичные локальные показатели эффективности. Возможным подходом в такой ситуации является применение разработанного в начале 1970 года американским математиком Томасом Саати метода анализа иерархий (Analytical Hierarchy Process) [11, 12].

Этот метод получил широкое распространение и активно применяется в различных отраслях для решения прикладных задач многокритериального оценивания на основе обработки субъективных экспертных суждений в форме парных сравнений. Метод применяется к различным по своей сущности и характеристикам системам и объектам. Так известны постановки и решения задач ранжирования и выбора возможных рыночных альтернатив (продуктов и услуг), проведения маркетинговых исследований, определения значимости совокупности факторов, стратегий развития, распределения ресурсов, и т.д.

В основе метода лежит процедура декомпозиции сложной проблемы – представление её в виде структурированного набора компонент или критериев, взаимосвязи между которыми формируются в иерархическом варианте представления. При этом вершиной иерархии является общая цель – желаемое состояние системы. Следующий уровень представляет собой детализацию общей цели в виде набора критериев, компонент или сил, оказывающих влияние на достижение обозначенного результата. На самом нижнем уровне иерархии представлены возможные альтернативы, степень приоритетности (значимости) которых требуется оценить.

В методе осуществляется процедура парного сравнения отдельных компонент иерархии между собой. Результаты экспертного оценивания представляются в виде набора положительных обратносимметричных матриц парных сравнений вида:

$$(9) \quad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

В (9) каждый элемент матрицы a_{ij} определяет субъективное мнение эксперта относительно значимости оцениваемой i -ой компоненты иерархии относительно j -ой.

Матрицы вида (9) составляются для сравнения важности элементов каждого нижестоящего уровня иерархии относительно вышестоящего. Оценки элементов иерархии производятся в соответствии с девятибалльной шкалой относительной важности, при этом для оценивания n компонент необходимо сформировать $n(n-1)/2$ суждений. На основе матриц экспертных суждений вычисляются локальные приоритеты (значимости факторов) по каждой отдельной матрице, и критерии качества экспертных оценок.

Локальные приоритеты формализуются в методе МАИ в виде нормализованного главного собственного вектора матрицы. Широко распространённым приближённым способом нахождения собственного вектора является подход, основанный на нахождении среднего геометрического. В соответствии с ним компонента собственного вектора i -ой строки находится по формуле:

$$(10) \quad b_i = \sqrt[n]{a_{i1} \times a_{i2} \times \dots \times a_{in}}.$$

Затем все компоненты b_n собственного вектора нормализуются на единицу путем их деления на сумму $\sum b_i$ – и эти значения определяют локальную значимость относительно оцениваемых элементов.

Критерии качества работы экспертов отыскиваются на основе расчёта показателей согласованности суждений для каждой матрицы парных сравнений по максимальному собственному значению матрицы.

Если качество суждений экспертов неприемлемо – оценки несогласованны между собой – то следует пересматривать суждения экспертов по соответствующей матрице.

Завершающим этапом метода является нахождение интегральных обобщённых оценок значимости альтернатив. Процедура свертывания локальных приоритетов заключается в нахождении взвешенных сумм по всем элементам одного уровня, учитывающих весовые коэффициенты (вектора приоритетов) вышестоящего уровня иерархии. Так в случае с трёхуровневой иерархией в соответствии с весовыми коэффициентами критериев второго уровня производится свёртка по каждой из альтернатив третьего уровня. Найдя, таким образом, значения глобальных приоритетов, делается окончательный вывод о сравнительной значимости оцениваемых альтернатив.

Рассмотрим применение метода МАИ для задачи оценивания конкурентоспособности. На одном из предприятий авиационного машиностроения Самарской области была проведена экспертная оценка факторов конкурентоспособности предприятия по следующей схеме:

1. Выбраны и согласованы показатели конкурентоспособности на основе существующей системы индикативных показателей.
2. Проведена структуризация показателей и факторов конкурентоспособности в виде иерархической структуры.
3. Проведено экспертное оценивание значимости факторов конкурентоспособности.
4. Реализована математическая обработка результатов анкетирования.
5. Найдены обобщённые значимости факторов конкурентоспособности, определены численные значения рангов и значимости полученных оценок.

Исходной информацией для проведения исследования явилась система оценочных показателей конкурентоспособности предприятия. На следующем этапе было произведено построение иерархии и формирование схемы проведения исследования – рисунок 6.

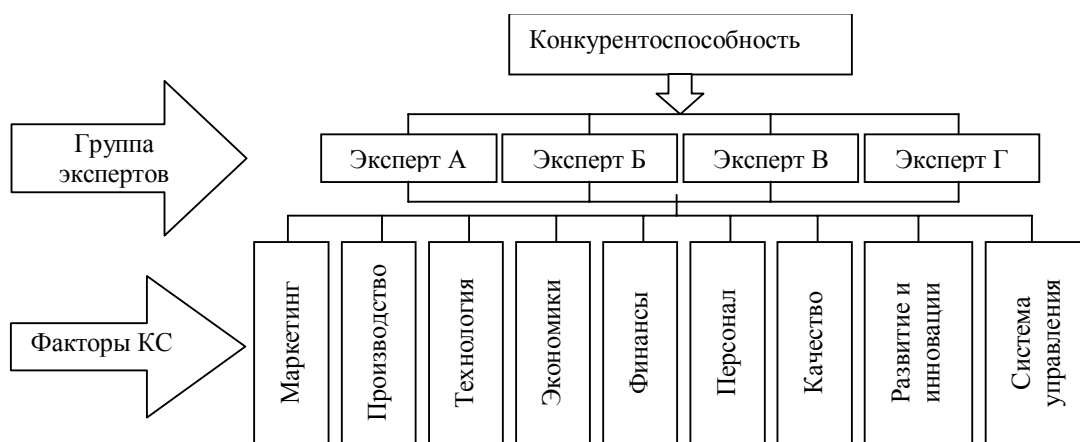


Рисунок 6 – Иерархия системы оценки факторов конкурентоспособности.

Было выделено девять факторов конкурентоспособности: маркетинг, производство, технология, экономика, финансы, персонал, качество, развитие и инновации и сфера управления. В качестве экспертов, осуществляющих ранжирование факторов конкурентоспособности были выбраны ведущие менеджеры предприятия. Ранжировка экспертов не проводилась, а степень значимости их суждений определялась на основе оценок качества – отношений согласованности. В случае если несогласованность превышала допустимый уровень, то при синтезе глобальных приоритетов использовались понижающие весовые коэффициенты.

Каждый эксперт провёл попарное сравнение значимости (степени влияния) девяти факторов на общую конкурентоспособность предприятия. Результаты агрегирования их заключений по описанной методике в глобальные критерии оценки представлены на диаграмме рисунка 7.

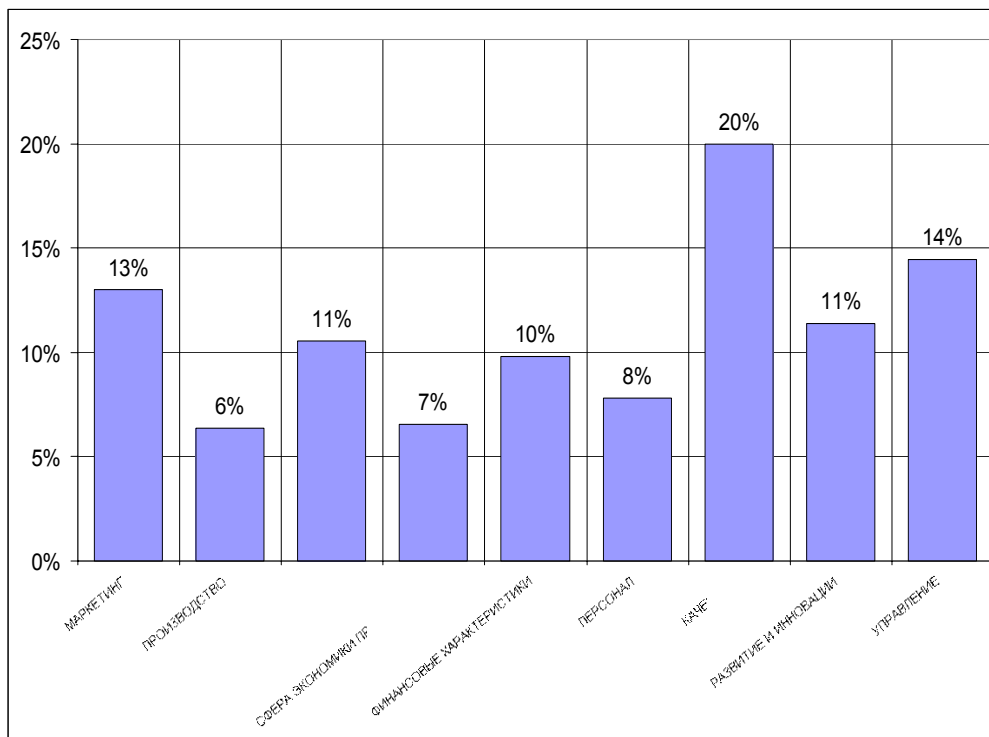


Рисунок 7 – Результаты оценки глобальных приоритетов по факторам конкурентоспособности.

Полученные обобщённые оценки факторов конкурентоспособности выделяют наиболее и наименее значимые показатели, дают возможность структурирования их для различных целей. Так, все факторы можно разделить на три группы приоритетов: «приоритетная» (оценки от 20% до 13%) – качество, управление, маркетинг; «средняя» (оценки на уровне 11%, 10%) – технология, инновации, финансы; «низшая» (оценки менее 8%) – персонал, экономика и производство. Качественной интерпретацией, полученных оценок, является возможное «благополучное» положение дел в группах среднего и низкого приоритетов и наиболее острой или актуальной ситуацией в приоритетной группе.

К ключевым особенностям рассмотренного метода МАИ относится то, что, во-первых, весовые коэффициенты оценочных критериев определяются в ходе реализации алгоритма парных сравнений, а не задаются априори. Во-вторых, рассматриваемый подход позволяет получать оценки на разных уровнях иерархии, что существенно облегчает задачу системного оценивания в случае большого количества альтернатив.

Отметим, что рассмотренный метод МАИ хорошо согласуется с методом многокритериального оценивания сравнительной эффективности – Data Envelopment Analysis. Так в работе [13] предлагается на предварительном этапе построения функциональной модели оценки сравнительной эффективности сначала оценивать значимость набора входных и выходных параметров методом МАИ и только потом использовать алгоритм расчёта в соответствии с DEA-методикой.

Заключение

В статье рассмотрены различные методы обобщённого, многокритериального оценивания, основанные как на детерминированных и статистических подходах, так и на методах экспертных суждений. Конструктивное применение этих подходов, основанных на построении функциональных и структурных моделей исследуемых объектов, позволяет выявлять и обосновывать наиболее эффективные направления совершенствования комплексной деятельности производственно-экономических систем.

Список литературы

- [1] Иванилов Ю.П., Лотов А.В. Математические модели в экономике. – М.: Наука, 1979. – 304с.
- [2] Терехов Л.Л. Производственные функции. – М.: Статистика, 1974. – 128с.
- [3] Самарский статистический ежегодник (К 150-летию Самарской губернии) / Самарский областной комитет государственной статистики. – Самара, 2001
- [4] Социально-экономическое положение Самарской области (1990 – 2001 гг.)/ Самарский областной комитет государственной статистики. – Самара, 1991 –2002.
- [5] Орлова Е.Ю. Структурно-параметрическая идентификация региональных технологических производств как объектов управления: Дис. ... канд. тех. наук: 05.13.07. – Самара, 1999. – 204с.
- [6] Гранберг А.Г. Моделирование социалистической экономики: Учеб. для студ. экон. вузов. – М.: Экономика, 1988.
- [7] Клейнер Г.Б. Производственные функции: теория, методы, применение. – М.: Финансы и статистика, 1986. – 239с.
- [8] Farrel M.J. The Measurement of Productive Efficiency // Journal of the Royal Statistical Society, Series A (General), Vol. 120, Part III, 1957, p. 253 – 281.
- [9] Banker R.D., Charnes A, Cooper W.W. Some Models for Estimating Technical and Scale Efficiency in Data Envelopment Analysis // Omega, The International Journal of Management Science, Vol. 30, No. 9, 1984, p. 1078 –1092.
- [10] Интрилигатор М. Математические методы оптимизации и экономическая теория. Пер.с англ. – М.: Прогресс, 1975.– 606с.
- [11] Саати Т. Принятие решений. Метод анализа иерархий. М.: Радио и связь, 1993. – 320 с.
- [12] Саати Т., Кернс К. Аналитическое планирование. Организация систем. М.:Радио и связь, 1991.–224с.
- [13] Changhoon Shin, Jaeyoung Song, Hyunchan Kim DEA/AHP Model for Measuring Restricted and Effective Efficiency with Expert's Knowledge // Informs&Korms – 1140 – Department of Logistic Engineering, Korea Maritime University – Seoul 2000.

СИНЕРГЕТИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ ОРГАНИЗАЦИОННЫХ СИСТЕМ

М.Б. Гузаиров, Р.А. Бадамшин, Б.Г. Ильясов, И.Б. Герасимова, Н.В. Хасанова

Уфимский государственный авиационный технический университет

450000, Уфа, ул. К.Маркса 12, Россия

ilyasov@tc.ugatu.ac.ru

тел: (8-3472) 23-78-35, факс: (8-3472) 22-29-18

Ключевые слова: синергетика, управление, организационная система, проект

Abstract

In the given paper a synergetic approach for analysis of realization efficiency of scientific projects in different scientific organizational structures is considered. Nonlinear models of implementation of one or two scientific projects for two and three performers with different relation variants between them are investigated.

Введение

Исследование процессов управления в организационных системах представляет как теоретический, так и практический интерес. Сложность исследования этих процессов обусловлена во-первых, наличием человека как элемента системы, поведение которого является источником различного рода неопределенностей, во-вторых, сложностью отношений (от дружелюбного до конфликтного) между субъектами, изменение которых ведет к изменению структуры организационной системы, в-третьих, неопределенностью уровня сложности научной проблемы, для решения которой и построена организационная система.

Исследование динамики организационных систем на основе нечетких когнитивных карт позволяет выявить влияние многих внешних факторов на эффективность функционирования системы, лучше понять проблему, раскрыть внутренние противоречия системы [1]. Однако линейные когнитивные модели не позволяют выявлять и анализировать те свойства системы, которые являются результатом нелинейного поведения элементов и нелинейного характера отношений между ними.

Исследования в нелинейных оргсистемах процессов управления как процессов самоорганизации целесообразно проводить на основе достижений синергетики, которая охватывает почти все современные отрасли знаний [2, 3, 4].

В данной статье рассматривается синергетический подход к анализу эффективности реализации научных проектов в различных научных организационных структурах. Исследуются нелинейные модели реализации одного и двух научных проектов при двух и трех исполнителях с различными вариантами отношений между ними.

1 Анализ взаимодействия участников проекта при его реализации

При изучении процессов управления в сложных нелинейных динамических системах является важным выявление роли положительных и отрицательных обратных связей в формировании различных форм движения системы, ее свойств, в том числе ее способности к достижению высокого уровня самоорганизации. Отличительная особенность изучаемого подхода к управлению в организационных системах заключается в том, что в них реализуется синергетический принцип управления, когда достижение цели осуществляется не за счет жесткого отслеживания состояния объекта управления, как это происходит в следящих системах, а за счет

такой организации взаимодействия участников проекта, которая и обеспечивает его эффективную реализацию.

При этом цель индивидуальной деятельности участников проекта заключается в такой организации их совместной деятельности, при которой удастся поддержать ритмичность и постоянный темп выполнения работ по научному проекту (НП), обеспечивающие реализацию проекта в намеченный срок. Если организация совместной деятельности участников проекта построена неэффективно, то наблюдается хаос в темпах реализации проекта, в темпах выполнения работ участниками проекта. Этот хаос может быть ликвидирован только за счет целенаправленного пересмотра отношений между участниками проекта и их индивидуального поведения, которые выступают в системе в качестве параметров порядка.

В работе рассмотрены несколько типов нелинейных динамических моделей организационного управления процессом реализации научного проекта. Все модели адекватно отражают качественную сторону взаимодействия участников проекта.

Первая модель описывает процессы взаимодействия научного руководителя (НР) и научного сотрудника (НС) при реализации НП и имеет следующий вид:

$$(1) \begin{cases} \dot{Y} = -a_1 Y + k_{12} XY - b_1 XZ + u_1, \\ \dot{X} = -a_2 X + k_{21} XY - b_2 YZ + u_2, \\ \dot{Z} = -cZ + XY. \end{cases}$$

Здесь Z – темп выполнения работ по НП, а Y и X соответственно темпы выполнения работ по НП научным руководителем и научным сотрудником. Параметры a_1, a_2 отражают факторы, стабилизирующие индивидуальную деятельность НР и НС. Коэффициенты k_{12}, k_{21} отражают взаимодействие между участниками проекта и образуют в системе положительную обратную связь. Параметры b_1, b_2 отражают факторы, стабилизирующие процесс поддержки темпов выполнения проекта в целом. Коэффициент c отражает влияние степени сложности проекта и трудоемкости его выполнения на стабилизацию темпов выполнения проекта. Переменные u_1, u_2 отражают вклад, вносимый участниками в поддержание темпа выполнения работ по научному проекту.

Предположим, что XY есть общее усилие интеллектуальной деятельности НР и НС, направленное, во-первых, на обеспечение темпа выполнения работ по НП, во-вторых, на оказание взаимопомощи друг другу. Причем в бесконфликтной ситуации эта помощь такова, что усиливает темпы выполнения работ НР и НС, а при конфликтной ситуации – снижает эти темпы. Кроме того, в моделях деятельности НР и НС учитывается их совместное восприятие процессов управления в виде стабилизирующей обратной связи по темпу выполнения работ по НП. При этом сила обратной связи для НР определяется НС, а для НС – его НР.

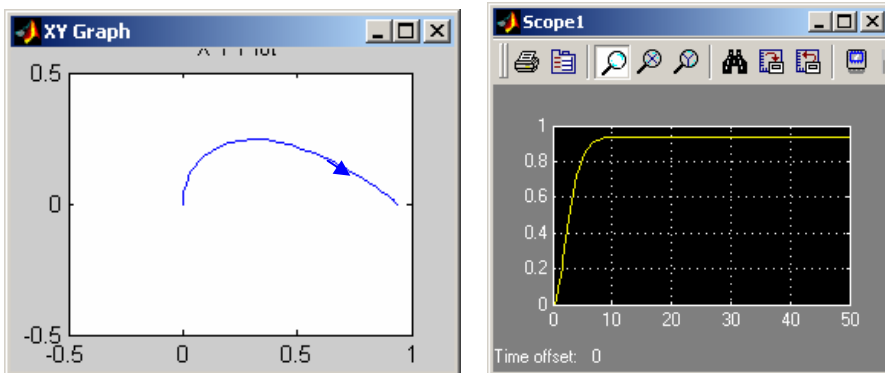


Рисунок 1 – Фазовый портрет и переходная функция в управляемой оргсистеме

На рисунках 1-3 представлены различные фазовые портреты и соответствующие им переходные процессы поведения системы в плоскости (Z, \dot{Z}) при различных значениях параметров системы. На рисунке 1 показан процесс поддержания постоянных темпов выполнения проекта при согласованном взаимодействии партнеров (НР и НС). Темп Z управляем по входным переменным u_1 и u_2 .

На рисунке 2 показан процесс, когда в системе возникает срыв типа «складки», переводящий систему с высокого уровня на более низкий уровень темпов выполнения проекта. Этот срыв происходит из-за низкой инерционности выполнения самого проекта как объекта управления и высокой инерционности деятельности исполнителей.

На некоторых режимах работы системы наблюдается возникновение автоколебательных процессов. На рисунке 3 показан срыв системы с переходом в автоколебательный режим.

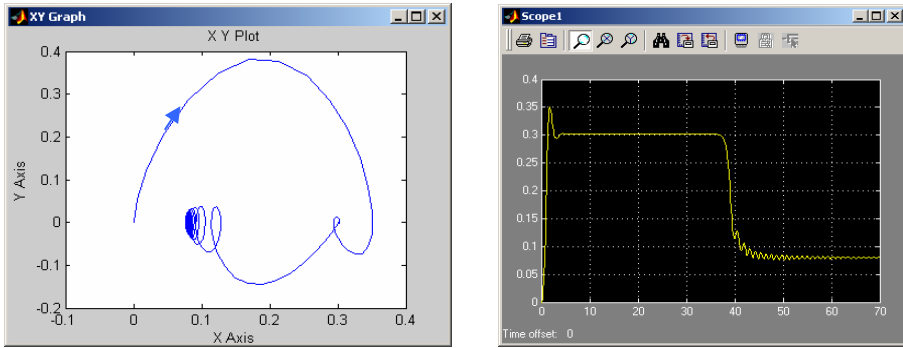


Рисунок 2 – Срыв в оргсистеме

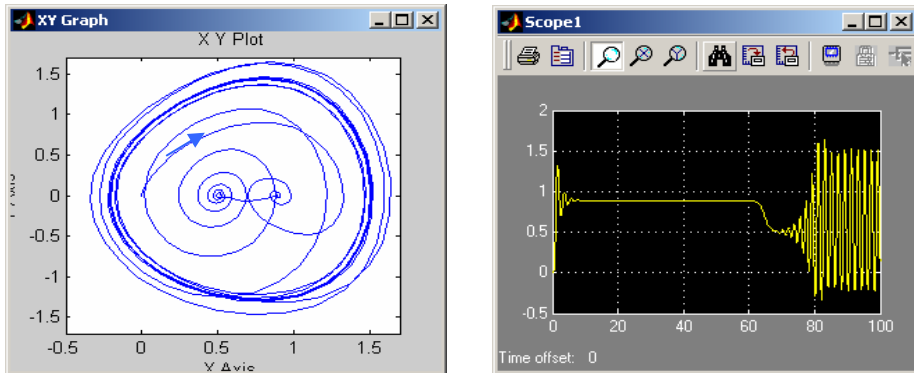


Рисунок 3 – Срыв с переходом в автоколебательный режим

Далее количество НС в модели было расширено до двух. Математическая модель такой оргсистемы имеет следующий вид

$$(2) \begin{cases} \dot{Y} = -a_0 Y + k_{01} X_1 Y + k_{02} X_2 Y - b_0 (X_1 + X_2) Z + u_0, \\ \dot{X}_1 = -a_1 X_1 + k_{10} Y X_1 - b_1 Y Z + u_1, \\ \dot{X}_2 = -a_2 X_2 + k_{20} Y X_2 - b_2 Y Z + u_2, \\ \dot{Z} = -c Z + (X_1 + X_2) Y. \end{cases}$$

В этой модели НР работает с каждым НС в отдельности, а НС между собой не контактируют. Если требуется учесть взаимодействие между НС, то в уравнения 2 и 3 системы (2) необ-

ходимо ввести члены $k_{12}X_1X_2$ и $k_{21}X_1X_2$ соответственно. Эффект от деятельности НС воспринимается суммарно, а при учете их взаимодействия в уравнении 4 системы (2) должен появиться член X_1X_2 . В этой модели также можно достичь режима поддержания постоянного темпа выполнения работ по НП. При ослаблении стабилизирующих связей здесь также возникает хаос в поддержании темпов выполнения проекта. Возникновение конфликтных ситуаций между НР и НС приводит к значительному (в 3÷5 раз) снижению темпов выполнения работ. В этой системе также могут появляться автоколебательные режимы (рис.4) и колебания типа «странного» аттрактора (рис. 5). При наличии запаздывания в каналах стабилизирующей обратной связи возникают более сложные формы хаотических движений.

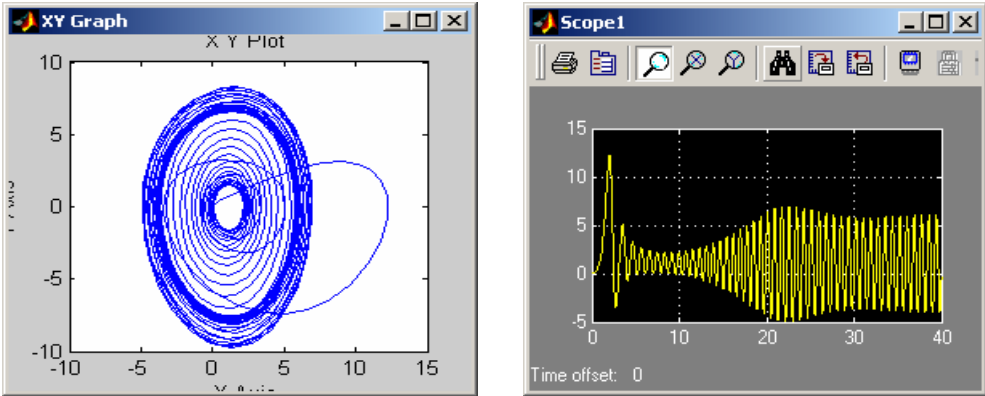


Рисунок 4 – Автоколебательный режим в оргсистеме

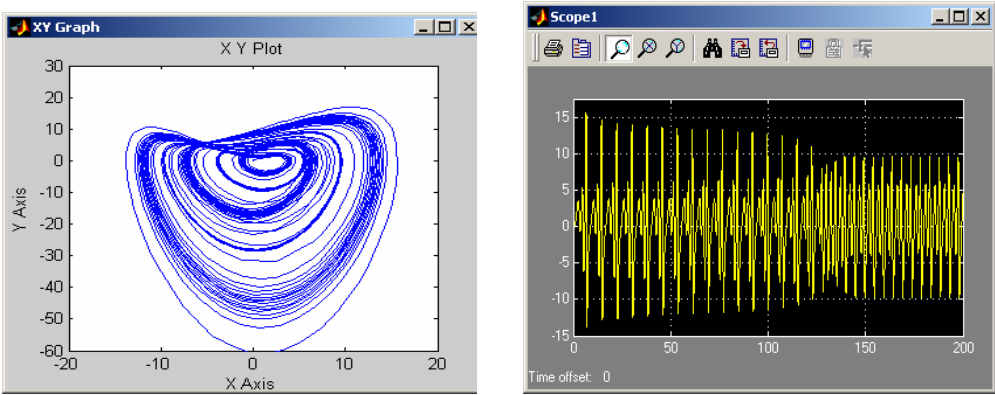


Рисунок 5 – «Странный» аттрактор

В качестве третьей модели была рассмотрена оргсистема, в которой реализуются два проекта при трех участниках проекта.

$$(3) \begin{cases} \dot{Y} = -a_0Y + k_{01}X_1Y + k_{02}X_2Y - b_{01}X_1(Z_1 + Z_2) - b_{02}X_2(Z_1 + Z_2) + u_0, \\ \dot{X}_1 = -a_1X_1 + k_{10}YX_1 - b_1YZ_1 + u_1, \\ \dot{X}_2 = -a_2X_2 + k_{20}YX_2 - b_2YZ_2 + u_2, \\ \dot{Z}_1 = -c_1Z_1 + (X_1 + X_2)Y, \\ \dot{Z}_2 = -c_2Z_2 + (X_1 + X_2)Y. \end{cases}$$

В этой модели не учтено взаимодействие между НС, но оба НС участвуют в реализации обоих проектов. В этой оргсистеме также можно за счет организации взаимодействия между участниками проекта обеспечить такой режим ее работы, при котором обеспечивается поддержка требуемых темпов выполнения работ по каждому проекту. На рисунке 6 показаны возможные типы хаоса в данной системе.

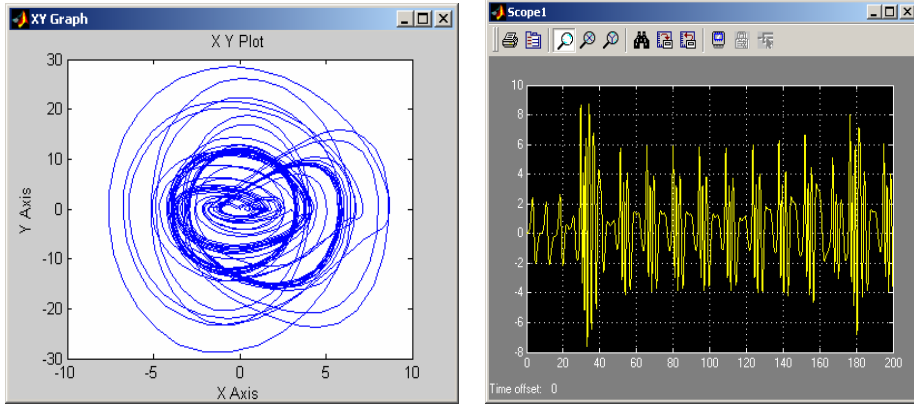


Рисунок 6 – Хаотические движения в системе при наличии чистого запаздывания

На рисунке 7 показана одна из схем моделирования оргсистемы.

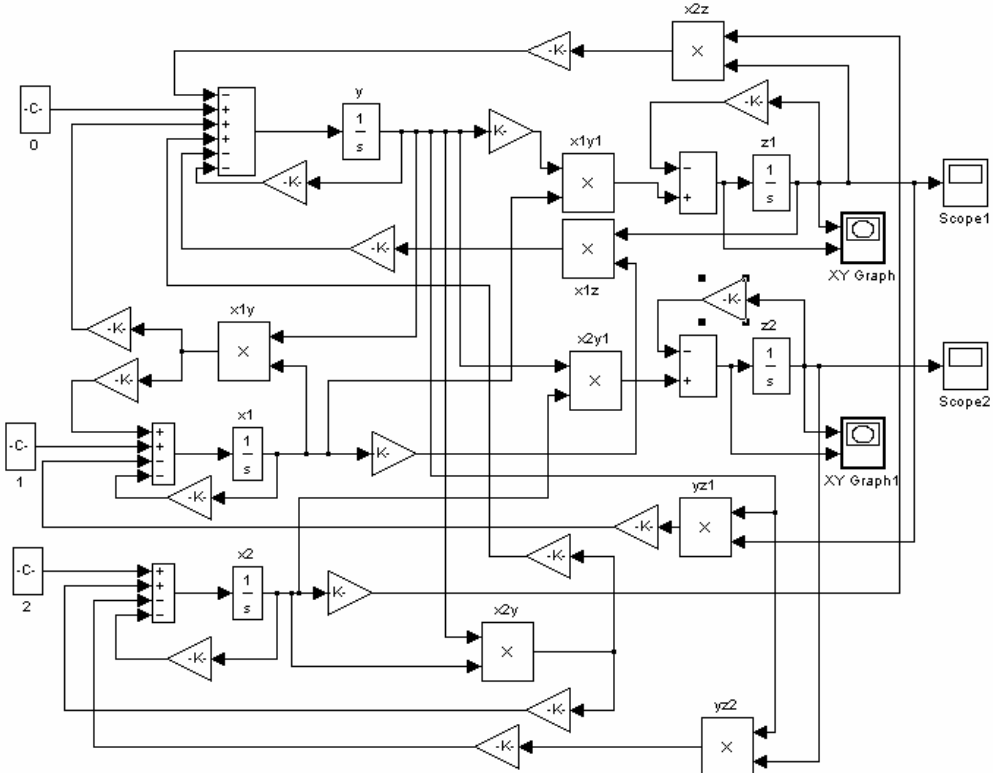


Рисунок 7 – Схема моделирования оргсистемы

Исследования, проведенные на данных моделях позволяют сделать следующие выводы.

- 1) За счет соответствующей организации отношений между участниками проекта можно достичь управляемых режимов поддержания желаемых темпов выполнения работ по научным проектам.
- 2) Для стабилизации темпов выполнения работ по НП должны быть соблюдены следующие условия:
 - инерционность индивидуальной деятельности участников проекта должна быть значительно ниже инерционности изменения темпов выполнения работ по проекту;
 - инерционность деятельности НС должна быть меньше, чем инерционность деятельности НР;
 - в системе должно отсутствовать чистое запаздывание в передаче информации и других ресурсов между участниками проекта;
 - в системе сила стабилизирующих связей должна быть больше силы стимулирующих (положительных) связей между НР и НС, направленной на взаимопомощь друг другу.
- 3) К дестабилизирующим факторам, вызывающим детерминированный хаос в поддержании темпов выполнения работ по НП, можно отнести следующие:
 - низкие значения стабилизирующих факторов: при малых значениях b_1, b_2, b_{01}, b_{02} система теряет статическую устойчивость;
 - большие значения стимулирующих связей между НР и НС, которые образуют в системе положительные обратные связи;
 - большая инерционность деятельности участников проекта по сравнению с инерционностью изменения темпов выполнения работ по проекту, что приводит к возникновению автоколебаний и даже к неустойчивости системы;
 - наличие чистого запаздывания в каналах обратной связи;
 - высокая чувствительность к воздействию факторов внешней среды.

Дальнейшие исследования направлены на усложнение нелинейной модели в сторону увеличения как количества элементов в системе, так и параметров, характеризующих состояние каждого из участников проекта.

Заключение

Синергетический подход позволяет исследовать в нелинейных динамических оргсистемах процессы управления, построенные на принципе самоорганизации систем, и выявить все возможные аттракторы и формы движения систем, а также получить условия, при которых система попадает в желаемый аттрактор.

Список литературы

- [1] Бадамшин Р.А., Ильясов Б.Г., Герасимова И.Б., Хасанова Н.В. Анализ эффективности решения научной проблемы в малых научных группах на основе динамических когнитивных карт // Проблемы управления и моделирования в сложных системах: Труды IV международной конференции (Самара, 17-23 июня 2002 г.). – Самара: СНЦ РАН, 2003. С. 182-189.
- [2] Хакен Г. Синергетика. – М.: Мир, 1980.
- [3] Пригожин И, Стенгерс И. Порядок из хаоса. – М.: Прогресс, 1986.
- [4] Современная прикладная теория управления: синергетический подход в теории управления / Под ред. А.А. Колесникова – Таганрог: Изд-во ТРТУ, 2000. – Ч.2. – 559 с.

САМООБУЧАЮЩИЕСЯ ПРЕДПРИЯТИЯ *

В.Б. Тарасов

Московский государственный технический университет им. Н.Э.Баумана
Россия, 107005, Москва, 2-я Бауманская ул., 5
tarasov@rk9.bmstu.ru
тел: +7 (095) 263-60-49, факс: +7 (095) 263-61-70

Ключевые слова: моделирование организационное, стратегия развития, обучение, поведение, адаптация, организация интеллектуальная, предприятие самообучающееся, обучаемость, формирование опыта, управление знаниями

Abstract

A new enterprise competition strategy related to permanent enterprise learning is suggested. The sources of learning enterprise concept are revealed. The relationships between learning and intelligent enterprises are considered. Some well-known learning procedures in psychology and artificial intelligence are discussed. Main functions and types of enterprise learning are studied. Formal approaches to selecting behavior strategies for learning enterprises are introduced. At last, a methodology of constructing learning enterprises is developed.

Введение

В современном бизнесе типичными источниками получения конкурентных преимуществ являются: новые технологии (в особенности, информационные и коммуникационные технологии); новые рынки (или новый сегмент рынка); изменения «правил игры» на рынке; новые запросы потребителей; наконец, новые структуры и стратегии поведения предприятий. В [1-4] были рассмотрены такие инновации организационных структур предприятий как горизонтальные, расширенные, фрактальные, холонические предприятия. Кроме того, в [1, 5] были подробно исследованы вопросы построения интеллектуальных предприятий. Однако, организационное поведение изменяется параллельно с модификацией структуры организации.

Цель настоящей работы заключается в исследовании новых стратегий организационного поведения предприятий, связанных с появлением обучающихся предприятий. При этом используется концепция «экобиологических организаций» как «социо-экономических организмов, живущих и развивающихся в сложной, динамической среде» [6]. Главная проблема создания обучающихся предприятий заключается в переходе от традиционных способов и средств индивидуального и группового обучения к обучению естественных и искусственных организаций.

Согласно М. Портеру, выделяются три классические конкурентные стратегии [7]: стратегия лидерства по издержкам, стратегия дифференциации и стратегия фокусирования. Стратегия лидерства по издержкам предполагает поддержание более низких издержек по сравнению с конкурентами. Основой стратегии дифференциации является придание товару отличительных свойств, которые важны для покупателя и отличают данный товар от предложений конкурентов. Хотя для этого обычно требуются более высокие издержки, успешная дифференциация позволяет компании добиться рентабельности благодаря тому, что рынок готов принять более высокую цену. Наконец, стратегия фокусирования означает специализацию на нуждах одного сегмента или конкретной группы покупателей без стремления охватить весь рынок. В ее основе лежит стремление удовлетворить потребности выбранного целевого сегмента луч-

* Работа выполнялась при поддержке РФФИ, проекты № 02-01-81037 и № 04-07-90440.

ше, чем конкуренты. Такая стратегия может опираться как на дифференциацию, так и на лидерство по издержкам, но только в рамках отдельного целевого сегмента.

В данной работе описана новая стратегия получения конкурентных преимуществ в эпоху экономики, основанной на знаниях [5]: *стратегия перманентного обучения предприятия*, связанная с непрерывным накоплением и интенсивным использованием *корпоративного опыта* (а в целом, интеллектуального капитала). Рассмотрены основные принципы и проблемы построения обучающихся и самообучающихся организаций.

1 Истоки концепции самообучающихся предприятий

В основе построения обучающихся предприятий лежат модели *организационного научения*, т.е. формирования нового опыта в процессе взаимодействия организации со средой (см. например, [8]), и предложенная в 1990 г. П. Сенджем в его книге «Пятая дисциплина» концепция *самообучающейся организации* [9, 10]. Последняя развивает подход системной динамики Дж. Форрестера и указывает пять основных взаимосвязанных дисциплин, выступающих как необходимые условия образования ядра обучающейся организации: 1) формирование профессионального опыта и личного мастерства; 2) построение и реорганизация внутренних (ментальных) моделей внешней среды и самой организации; 3) создание общей картины организационного развития, построение «образа потребного будущего» предприятия; 4) осуществление группового обучения; 5) системное мышление (пятая и объединяющая дисциплина).

Согласно Г. Саймону, под научением понимается любое изменение в системе, приводящее к улучшению решения задачи при ее повторном предъявлении или к решению другой задачи на основе все тех же знаний. Организационное научение направлено на повышение эффективности деятельности предприятия, его приспособление к среде функционирования. Оно предполагает активное овладение новыми стратегиями и способами деятельности.

Каковы же главные характеристики самообучающихся предприятий? На наш взгляд, к их числу относятся:

- способность организационной структуры к автономному функционированию в быстро меняющейся среде благодаря корпоративному опыту, оперативному восприятию изменений и использованию прогностических стратегий;
- формирование, накопление, развитие, обобщение и продуктивное использование своего и чужого опыта; наличие специальных средств, обеспечивающих «информационную прозрачность» и циркуляцию знаний между подразделениями предприятия, а также между предприятием и его партнерами (получение обратной связи от клиентов и поставщиков);
- проведение процедур «эталонного тестирования» или «сопоставления с образцом» (benchmarking), т.е. сравнительная оценка продуктов, услуг, методов и средств предприятия по отношению к наиболее сильным конкурентам или мировым лидерам в данной области; построение на этой основе «образа будущего предприятия»;
- приоритетное использование стратегий развития человеческих ресурсов, стимулирование познавательной мотивации и создание климата, благоприятствующего самообучению, групповому обучению и развитию творческого потенциала сотрудников;
- осуществление социального поведения, т.е. формирование у сотрудников общих ценностей, единых интересов и стремлений, совместных намерений и взаимных обязательств, связанных с дальнейшей судьбой предприятия; распространение своей организационной культуры на партнеров (поставщиков, подрядчиков, клиентов);
- вхождение в различные организационные объединения и альянсы, умение извлекать из этого экономическую выгоду.

По информации известных американских консультантов в области стратегического менеджмента и бизнеса Дж.К. Беннета и М.Дж. О'Брайена (см. [9, 11, 12]), еще в 90-е годы XX века объявили о своем стремлении стать самообучающимися предприятиями и сделали кон-

клетные шаги на этом пути такие крупные транснациональные компании как Хегох, Kodak, General Electric, Motorola, Wal-Mart. Так президент компании Моторола Б. Гэлвин утверждает, что основными способами поддержания конкурентоспособности фирмы в экономических сражениях будущего видятся: а) оперативное распознавание и быстрое удовлетворение потребностей клиентов; б) адаптация к постоянно меняющейся среде; в) повышение креативности, творческого потенциала компании. Соответственно, компанией Моторола создан университет, где уже почти 10 лет проводится эксперимент по непрерывному образованию сотрудников в течение их жизни, на которое тратится 4-5% средств от общего дохода фирмы. Главные цели обучения таковы: 1) объяснить каждому сотруднику взаимосвязи между основными бизнес-процессами компании, так чтобы она работала как единый организм; 2) благодаря целенаправленному формированию и накоплению интеллектуального капитала (в частности, корпоративных знаний) всячески развивать активность и автономность компании в интересах быстрого самостоятельного освоения ею новых рынков и разработки новых технологий.

Основные принципы обеспечения эффективности организационного обучения предприятий таковы:

- обучаться быстрее, чем конкуренты;
- обучаться быстрее, чем изменяется внешняя среда (*принцип опережающего обучения*);
- обучаться внутри организации (сотрудникам друг у друга и в рамках рабочих групп);
- обучаться за пределами организации (у поставщиков и потребителей)
- обучаться по вертикали (от вершины до основания организации);
- прогнозировать будущее, создавать сценарии и обучаться по ним;
- признавать ошибки как неотъемлемый атрибут обучения и применять обучение в реальной деятельности.

2 О соотношениях между самообучающимися и интеллектуальными предприятиями

С одной стороны, для раскрытия соотношений между интеллектуальными и самообучающимися предприятиями приведем определение понятия «организационный интеллект» как набора способностей организации (см. также [1, 5, 13]). По нашему мнению, термин *организационный интеллект* охватывает способности организации:

- направленно *коэволюционировать* вместе со своей средой, т.е. успешно *адаптироваться* к изменениям внешней среды, адекватно, реагировать на любую, в особенности, новую ситуацию путем надлежащих корректировок *поведения*;
- *обучаться* и *самообучаться*, т.е. *приобретать*, *накапливать* и *использовать* опыт (знания, умения, навыки); *понимать* связи между фактами действительности для выработки действий, ведущих к достижению поставленной цели;
- формировать новые *коллективные (сетевые) структуры*, способствующие повышению живучести организации;
- строить адекватные *модели внешней и внутренней среды*, включая модели своего текущего и будущих состояний, в особенности, модель *потребного будущего*;
- осуществлять *самоуправление* своим *поведением* и *развитием*;
- эффективно *мобилизовать ресурсы*, необходимые для обеспечения эффективности и конкурентоспособности.

Таким образом, организационный интеллект (в широком смысле) рассматривается как многомерная, многоуровневая и динамическая система, представляющая собой совокупность структур, обеспечивающих выполнение определенных функций. Соответственно, организационное самообучение представляет собой отдельный аспект организационного интеллекта, а самообучающиеся предприятия могут трактоваться как специальный случай интеллектуальных предприятий.

С другой стороны, природа организационного интеллекта предприятия раскрывается через процедуры его формирования. Интеллект выступает как средство построения «базового поведенческого репертуара» предприятия, который строится благодаря определенным обучающим процедурам.

В целом, организационное поведение предприятия предполагает организационные изменения, обусловленные наличием следующих основных составляющих: *организация*, ее *внешняя и внутренняя среда*, *ресурсы* (капитал), *взаимодействие* организации со средой, предполагающее существование средств восприятия и первичной переработки информации из среды (*рецепторов*) и средств воздействия на среду (*эффекторов*), *память*.

Организационное поведение есть активная форма приспособления предприятия к среде; для того, чтобы оно было адекватно среде, необходимо, чтобы среда отражалась организацией. Память организации означает ее способность фиксировать и хранить информацию о своем состоянии и состоянии среды; по сути, она характеризует ее степень зависимости от прошлого. Память, включающая некоторый конкретный механизм кодирования, хранения и передачи информации, необходима для формирования избирательной реакции организации на различные воздействия.

Внешняя среда организации подразделяется на макросреду и микросреду. Макросреда определяется экономическими, социальными, политическими, технологическими факторами. В свою очередь, микросреда охватывает государственные экономические структуры, клиентов, поставщиков, посредников, конкурентов. Внутренняя среда включает сферы производства, маркетинга, финансов, управления персоналом и пр.

Обучение и самообучение выступают как важнейшие характеристики *комплексного поведения* организации (достигшей стадии организмического развития [14]), которое включает в себя такие формы (единицы) поведения как *рефлекс*, *инстинкт*, собственно *интеллект* (в узком смысле). *Реактивное поведение* организации, опирающееся на рефлекс, связано с выработкой своевременных ответных реакций на воздействия среды (например, заказы товаров или действия конкурентов). *Инстинктивное поведение* организации связано с ее приспособлением к определенным жизненным условиям, например, адаптацией к заданному состоянию среды, что предполагает координацию различных подсистем организации. Инстинкты считаются генетически закрепленными формами поведения, общими для данного класса организаций; они образуются в процессе эволюционного развития и практически не зависят от опыта конкретной организации.

Ситуационное поведение организации тесно связано с управлением жизненным циклом ее продукции (товаров) и предполагает умение добиваться поставленных целей. В частности, построение *адаптивного поведения* организации означает установление принципов функционирования в динамической, плохо определенной и быстро меняющейся среде, а реализация *подражательного поведения* организации связана с формированием новых способов поведения путем имитации действий другой организации. *Импульсивное поведение* организации означает быстрое выполнение действий под влиянием внешних обстоятельств.

Наконец, *интеллектуальное поведение* организации неотделимо от ее способностей отражать и прогнозировать потребности рынка, осознавать и контролировать свои мотивы, управлять знаниями. Оно также предполагает формирование стратегий организационного развития (например, стратегий получения конкурентных преимуществ, стратегий объединения и т.п.) и механизмов выбора вариантов взаимодействия со средой. В рамках интеллектуального поведения организаций могут рассматриваться социальное, нормативное, ролевое и другие формы поведения.

Различные виды и модели поведения живых и искусственных систем описаны в [15-17].

Таким образом, в поведенческом аспекте интеллектуальные предприятия могут пониматься как одна из высших форм самообучающихся предприятий. Примерами других, более низших форм самообучающихся организаций, служат реактивные, инстинктивные, импульсивные организации.

3 Основные функции и виды обучения и самообучения предприятий

В интересах построения общей модели самообучения предприятий первоначально рассмотрим основные функции самообучения и соответствующие виды организационного поведения (таблица 1). Функциями самообучения предприятия являются: формирование опыта; отражение внешней среды и самого предприятия; саморегулирование на основе интенций (мотивации); определение стратегий поведения или путей развития.

Главной функцией самообучения предприятий является формирование и управление *корпоративным опытом*, в который включаются корпоративные знания, навыки и умения. Соответственно, процесс самообучения организации предполагает построение жизненного цикла ее знаний, развитие организационных навыков и выработку умений (*ноу-хау*). Ранее в [5] нами была предложена X-образная схема интенсивной организации корпоративного опыта (формирования организационных умений). Здесь результативность самообучения означает возрастание объема и «действия» корпоративного опыта со временем.

Таблица 1 – Взаимосвязь между функциями самообучения и видами поведения.

Функция обучения \ Виды поведения	ОПЫТ	ПОЗНАНИЕ	ИНТЕНЦИИ	РАЗВИТИЕ
ИНТЕЛЛЕКТУАЛЬНОЕ	Знания	Потребности	Мотивы	Стратегии
СИТУАЦИОННОЕ	Умения	Продукты	Цели	Тактики
РЕФЛЕКТОРНОЕ	Навыки	Признаки	Задачи	Реакции

Основными видами обучения являются привыкание, классическое и оперантное обусловливание, комплексное научение. Простейший вид обучения предприятия – *привыкание* – представляет собой адаптацию предприятия к повторяющимся событиям (например, к аудиторским проверкам). *Классическое обусловливание* означает установление связей между событиями; хорошим примером выработки подобного *условного рефлекса* на предприятиях может служить организация поставок по схеме «деньги–товар». *Оперантное обусловливание* предполагает приобретение нового опыта и выработку новых способов поведения для достижения цели. Например, опыт сотрудничества с бывшим конкурентом приводит к возникновению поведения, опирающемуся на стратегию организационной (а не конкурентной) рациональности. Важным частным случаем оперантного обусловливания является *обучение с подкреплением*, когда происходит поощрение или наказание обучаемого в зависимости от результатов его действий. Здесь речь идет об использовании объектов или событий, значимых для организации, при ее обучении, когда должны запоминаться соответствия между ситуациями и действиями. *Импринтинг* означает фиксацию и запоминание объектов, находящихся в ближайшем окружении, и их главных признаков. В свою очередь, при *ассоциативном обучении* запоминаются связи между различными ситуациями. Наконец, *комплексное обучение* предполагает не только возникновение новых связей между событиями или появление новых форм поведения, но и развитие новых стратегий деятельности предприятия, а также приобретение новых знаний о среде и самой организации. Типичными примерами служат: *перенос* – использование ранее приобретенного опыта в новых ситуациях для последующего формирования опыта; *инсайт* – моментальное озарение.

Теперь сформулируем отличительные особенности различных уровней самообучения предприятий, используя таблицу 1. Здесь самообучение на уровне *условных рефлексов*, направленное на выработку ключевых *навыков предприятия*, предполагает определение основных *признаков ситуации*, выделение соответствующих *задач* и отработку *адекватных реакций*. Самообучение в русле *ситуационного поведения*, требующего наличия у предприятия развитых *умений*, строится вокруг *создания (моделей) продуктов и услуг* предприятия, анализа и генера-

ции *целей* предприятия, исследования *способов* решения возникающих *задач*. Наконец, самообучение *интеллектуальному поведению*, связанному с превращением *знаний* предприятия в преимущества над конкурентами, подразумевает осознание важности мотивационной сферы и развитие интенциональной основы деятельности, изучение вопросов *стратегического планирования* и *моделирования будущего* предприятия, его *самосохранения* и *самореорганизации*.

Нельзя не отметить, что в реализации «сквозных процессов обучения» на предприятиях центральное место занимают процессы *общения*. К числу основных функций общения на предприятии (и между предприятиями), определяющих различные способы *обучающих воздействий*, относятся: 1) *информационная (когнитивная)*, в частности, информирование, обоснование, объяснение; 2) *регулятивная*, например, убеждение, групповое давление, принуждение; 3) *аффективная*, например, внушение, заражение. Указанные обучающие воздействия направлены на различные компоненты ментальных моделей в организации: так, например, убеждение апеллирует к коллективному сознанию, а внушение – к коллективному бессознательному. Вариант классификации способов обучающих воздействий приведен в таблице 2.

Таблица 2 – Классификация способов обучающих воздействий в организациях.

Сфера приложения Выполняемая функция	Ментальная модель организации	
	КОЛЛЕКТИВНОЕ СОЗНАНИЕ	КОЛЛЕКТИВНОЕ БЕССОЗНАТЕЛЬНОЕ
КОММУНИКАТИВНО- КОГНИТИВНАЯ	Объяснение	Внушение
КОММУНИКАТИВНО- РЕГУЛЯТИВНАЯ	Убеждение Давление	Внушение Заражение

Можно также проследить взаимосвязи между видами обучающих воздействий и характеристиками *обучаемости* на предприятиях. Например, эффективность *объяснения* зависит от достигнутого уровня *понимания*, а возможность *убеждения* неразрывно связана с поиском *компромисса* и *достижением согласия*. Среди прочих важных характеристик обучаемости предприятий следует указать: *подражание* – следование предприятия избранному примеру; *конформизм* – пассивную адаптацию предприятия к воздействиям среды; *внушаемость* – готовность организации подвергнуться и подчиниться обучающему воздействию. Взаимосвязи между видами обучающих воздействий и характеристиками обучаемости показаны в таблице 3.

Таблица 3 – Взаимосвязи между видами обучающих воздействий и характеристиками обучаемости организаций.

ВИДЫ ОБУЧАЮЩИХ ВОЗДЕЙСТВИЙ	ХАРАКТЕРИСТИКИ ОБУЧАЕМОСТИ ОРГАНИЗАЦИЙ
Объяснение	Понимание
Убеждение	Согласие (компромиссность)
Заражение	Подражание
Давление среды (конформирование)	Конформизм
Внушение	Внушаемость

4 Определение стратегий поведения самообучающихся предприятий

Остановимся на проблеме формирования полного репертуара стратегий поведения самообучающихся предприятий и выбора эффективной стратегии. Изложим методику выбора, опирающуюся на использование следующих оппозиционных шкал критериев: «активность – пассивность», «индивидуальные действия – коллективные действия», «собственные цели пред-

приятия – цели партнеров» [13]. В первом приближении ограничимся двузначными полярными оценками по критериям, которые будем обозначать + (да) и – (нет). Тогда рассмотрение стратегий организационного поведения можно представить с помощью булевой решетки вида $2^3 = 8$. Возможные варианты поведения приведены в таблице 4.

Таблица 4 – Основные стратегии поведения предприятий.

ПРОЦЕСС ВЗАИМОДЕЙСТВИЯ	КРИТЕРИИ ВЗАИМОДЕЙСТВИЯ (НА ПОЛЯРНЫХ ШКАЛАХ)		
	Активность – Пассивность	Совместные действия – Индивидуальные действия	Общие цели – Свои цели
Стратегии поведения предприятий			
КООПЕРАЦИЯ	+	+	+
СОДЕЙСТВИЕ	+	+	–
КОМПРОМИСС	+	–	+
КОНКУРЕНЦИЯ	+	–	–
КОНФОРМИЗМ	–	+	+
СОЛИДАРНОСТЬ	–	+	–
ПРИСПОСОБЛЕНИЕ	–	–	+
УКЛОНЕНИЕ	–	–	–

Активность предприятий, взаимодействующих на рынке, приводит к возникновению отношений *кооперации* (сотрудничества) или *конкуренции* (соперничества). Первые возможны тогда, когда предприятия способны к формированию общей (совместной) цели и предпочитают совместную работу индивидуальной. В противном случае, зарождаются конкурентные отношения. Когда предприятие ориентировано лишь на свои интересы и не желает компромиссов, то оно руководствуется принципом *конкурентной рациональности*, и возникший конфликт выливается в конфронтацию до полной победы одного из конкурентов. Наоборот, когда предприятие предпочитает коллективные действия, то на первый план выходит стратегия *содействия* на рынке. Когда же оно способно подчинить свои индивидуальные интересы общим интересам партнеров, вполне естественной становится принятие им *компромиссной стратегии* поведения, например, связанной с формированием общей цели и коллективных стратегий ее достижения.

К пассивным вариантам поведения предприятий на рынке относятся *конформизм*, *солидарность* и *приспособление*. Здесь предприятие стремится адаптироваться к ситуации взаимодействия, отказываясь от своих целей и принимая требования другого предприятия. Стратегия солидарности характеризует «вынужденное содействие», когда совместные действия навязываются извне, а приспособление означает принятие чужих целей. Наконец, стратегия *уклонения* означает полный отказ от взаимодействия.

Аналогично можно конкретизировать различные варианты кооперации и конкуренции предприятий. Для этого следует взять такие базовые критерии как совместимость целей, потребность в чужом опыте, наличие общих заказов [1] (таблица 5). Тогда простое сотрудничество обусловлено необходимостью интеграции опыта предприятий-партнеров, а координируемое сотрудничество предполагает рациональное совместное использование распределенных ресурсов. В свою очередь, соперничество может быть как индивидуальным, так и коллективным. Наиболее продуктивным вариантом соперничества видится коллективное соперничество за ресурсы, когда конкурирующие группы способствуют ликвидации монополий и расширению рынка.

Конечно, все эти базовые стратегии и варианты поведения предприятий являются идеализацией и задают всего лишь исходные ориентиры организационного развития. Диапазон реального взаимодействия предприятий значительно шире и не может быть описан в рамках булевой

шкалы. Здесь целесообразно перейти к более общим алгебраическим конструкциям, например, алгебрам Клини с центральным элементом.

Таблица 5 – Типы кооперации и конкуренции.

Критерии ВАРИАНТ ПОВЕДЕНИЯ	Совместимость целей (да +, нет –)	Потребность в чужом опыте (да +, нет –)	Наличие Общих заказов (да +, нет –)
Координируемое сотрудничество	+	+	+
Простое сотрудничество	+	+	–
Непродуктивное сотрудничество	+	–	+
Независимость	+	–	–
Коллективное соперничество за заказы	–	+	+
Чистое коллективное соперничество	–	+	–
Индивидуальное соперничество за заказы	–	–	+
Чистое индивидуальное соперничество	–	–	–

5 Подходы к обучению в искусственном интеллекте и их использование на самообучающихся предприятиях

Основными направлениями развития искусственного интеллекта являются символьное, коннекционистское и бионическое (см. [13]). Соответственно, основные подходы к обучению в ИИ также делятся на три класса: символьные, нейросетевые и биосоциальные. К числу символьных методов относятся алгоритмы индуктивного обучения (поиск в пространстве вариантов, алгоритм ID3 и др.), алгоритмы обучения на основе сходства, алгоритмы обучения по аналогии, алгоритмы обучения по примерам, алгоритмы обучения с подкреплением. Под обучением нейронной сети понимается настройка архитектуры сети и весов связей для эффективного выполнения специальной задачи. Алгоритмы нейросетевого обучения подразделяются на алгоритмы обучения с учителем (т.е. на базе ошибок) и алгоритмы обучения без учителя. Первые обычно опираются на метод среднеквадратической ошибки или метод обратного распространения, а вторые используют правило Хебба, обучение методов соревнования и т.п. Среди биосоциальных подходов следует указать: генетические алгоритмы, системы классификации и генетического программирования, подходы искусственной жизни.

По нашему мнению, большой интерес для создания самообучающихся организаций представляют модели искусственной жизни, в особенности, модели возникновения целенаправленного адаптивного поведения в конкурентной среде [13, 17]. В частности, изучение взаимодействия мотивов и определение ситуационных правил поведения позволяют на модельных примерах обосновать выбор стратегии поведения обучающегося предприятия [18]. Проведенные эксперименты позволили раскрыть особенности и условия применения силовых, пионерских, нишевых и приспособительных стратегий предприятия.

6 Мероприятия по формированию обучающихся организаций

Предлагается следующая методика построения обучающихся предприятий.

- 1) Проявляется инициатива сверху: предприятие объявляет о своем намерении стать обучающейся организацией. Составляется список основных мероприятий для реализации данного намерения.

- 2) Предприятие создает внутри себя *климат информационной прозрачности* и поощряет обмен *знаниями и умениями* между сотрудниками и подразделениями. Для претворения в жизнь стратегии непрерывного образования сотрудников при предприятии создается корпоративный университет.
- 3) Активно используются *передовые информационные и коммуникационные технологии* для обеспечения более интенсивного *профессионального общения* между сотрудниками.
- 4) Строится и тиражируется модель *потребного будущего предприятия*, возникают общие представления о том, чему надо *научиться*, чтобы достичь этого будущего.
- 5) У сотрудников предприятия формируются *познавательная мотивация* и общий *интерес к будущему* предприятия, что способствует развитию единой *корпоративной культуры*.
- 6) *Системы оценки и вознаграждения* сотрудников строятся по принципу *наибольшего благоприятствования непрерывному обучению*.
- 7) *Руководство предприятия* вменяет в обязанность подразделениям *коллективное обучение* и постоянное *накопление опыта*.
- 8) *Менеджеры* принимают самостоятельные решения, которые способствуют эффективному обучению и развитию сотрудников.
- 9) Структура *профессиональных (должностных) обязанностей сотрудников* становится более гибкой в интересах обеспечения *оперативной адаптации* к среде (ротация и универсализация сотрудников, создание автономных междисциплинарных рабочих групп, и т.п.).
- 10) Работа в составе междисциплинарной рабочей группы понимается как важный *способ обучения*, в особенности, обучения профессиональной деятельности, обучения на примерах и обучения на ошибках.
- 11) Реализуются специальные процедуры *коллективного обучения* на предприятии, такие как групповое решение задач, методы технического творчества, и др.
- 12) Формируются новые *стратегии и способы организационного поведения* и, соответственно, новые *организационные структуры* предприятия.

Заключение

Самообучающиеся предприятия воплощают в жизнь новую стратегию организационного поведения и развития, связанную с реализацией различных функций и видов самообучения (и обучения) на индивидуальном, коллективном и организационном уровнях. Необходимым условием их построения является разработка следующих ключевых теоретических направлений:

- общая теория и модели обучения (охватывающая проблемы самообучения естественных и искусственных систем на индивидуальном, коллективном и организационном уровнях);
- теория самоорганизующихся, самообучающихся и развивающихся систем;
- инжиниринг организационных решений и действий в области самообучающихся предприятий;
- когнитивное моделирование предприятий, направленное на формирование и визуализацию их потребного будущего

На наш взгляд, общая теория обучения должна содержать такие разделы как: сквозное формирование опыта (сотрудника, отдела, предприятия, отрасли); овладение необходимыми в поведении стратегиями, тактиками, реакциями; теория когнитивно-регулятивной координации, универсальный организационный эволюционизм как единая теория развития.

Список литературы

- [1] Тарасов В.Б. Новые стратегии реорганизации и автоматизации предприятий: на пути к интеллектуальным предприятиям // Новости искусственного интеллекта. 1996. №4. С.40-84.
- [2] Варнеке Х.-Ю. Революция в предпринимательской культуре. Фрактальное предприятие: Пер. с нем. – М.: МАИК Наука/Интерпериодика, 1999.

- [3] Витих В.А., Лукс А.А., Мажаров Л.Г. Голонические производственные системы (обзор) // Проблемы управления и моделирования в сложных системах: Труды Международной конференции (Самара, июнь 1999). – Самара: СНЦ РАН, 1999. С. 380-385.
- [4] Ostroff F. The Horizontal Organization. – New York: Oxford University Press, 1999.
- [5] Тарасов В.Б. Интеллектуальные предприятия и управление знаниями: на пути к синергетическому искусственному интеллекту // Проблемы управления и моделирования в сложных системах: Труды IV международной конференции (Самара, июнь 2002). – Самара: СНЦ РАН, 2002. С. 180-191.
- [6] Тарасов В.Б. Искусственная жизнь – новый подход к моделированию и реорганизации предприятий // Проблемы управления и моделирования в сложных системах. Труды III международной конференции (Самара, сентябрь 2001). – Самара: СНЦ РАН, 2001. С. 166-176.
- [7] Портер М. Международная конкуренция: Пер. с англ. – М.: МО, 1993.
- [8] Арджирис К. Организационное научение: Пер. с англ. – М.: ИНФРА-М, 2004.
- [9] Сендж П. Пятая дисциплина. Искусство и практика самообучающейся организации: Пер. с англ. – М.: ЗАО «Олимп-бизнес», 1999.
- [10] Мильнер Б.З. Теория организаций (4-е изд). – М.: ИНФРА-М, 2004.
- [11] Garvin D.A. Building a Learning Organization // Harvard Business Review. 1993. Vol.73. P. 78-91.
- [12] Kelly K., Burgows P. Motorola: Training for the Millenium // Business Week. 28 March 1994.
- [13] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям. – М.: Эдиториал УРСС, 2002.
- [14] Моисеев Н.Н. Современный рационализм. – М.: МГВП КОКС, 1995.
- [15] Гаазе-Рапопорт М.Г., Поспелов Д.А. От амёбы до робота: модели поведения. – М.: Наука, 1987.
- [16] Кузнецов Н.А., Любецкий В.А., Чернавский А.В. К вопросу о понятии информационного взаимодействия. 2. Доречевой интеллект // Проблемы управления и моделирования в сложных системах. Труды III международной конференции (Самара, сентябрь 2001). – Самара: СНЦ РАН, 2001. С. 25-42.
- [17] Редько В.Г. Эволюционная кибернетика. – М.: Наука, 2003.
- [18] Картежников Д.И., Субботин А.А., Тарасов В.Б. Подходы искусственной жизни в моделировании взаимодействий и стратегий поведения предприятий // Научная сессия МИФИ-2004. Сборник научных трудов. Т.3. Интеллектуальные системы и технологии (Москва, январь 2004). – М.: МИФИ, 2004. С. 82-83.

МОДЕЛИ ИСКУССТВЕННОЙ ЖИЗНИ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ ЗАДАЧ В СЛОЖНЫХ ДИСКРЕТНЫХ СИСТЕМАХ

В.В. Емельянов, П.В. Афонин

МГТУ им. Н.Э. Баумана
107005, Москва 2-я Бауманская, 5
evv@rk9.bmstu.ru
тел: (095) 263-66-39

Ключевые слова: оптимизация, генетический оператор, популяция, модель, имитация, искусственная жизнь

Abstract

In the majority of researches the multi-agent system is considered without the account of its evolution as populations of the agents. The agents cooperate with each other, decide the subtasks, strengthen actions of other agents by the resources etc., but not "live" in a population, though adapt the characteristics during interaction with other agents. However if to carry out parallels between community of such intellectual agents and biological systems, the necessity of improvement by each agent of the functioning is visible during existence of a population, accumulation by this population of vital experience and as a consequence increase of volume of knowledge by each agent. Here already each agent (or group of the agents), as in the centralized control system, decides completely some task, but makes it in parallel with other agents (groups). Therefore population of such agents in the evolution receives at once so much decisions of a task, facing to them, how many is present in it of the intellectual agents (groups of the agents). This basic difference of the multi-agent system and systems of artificial life (in a broad sense of this concept) from other evolutionary algorithms.

Введение

В большинстве исследований многоагентная система (МАС) рассматривается без учета ее эволюции как популяции агентов. Агенты взаимодействуют друг с другом, решают свои подзадачи, усиливают действия других агентов своими ресурсами и т.п., но не «живут» в популяции, хотя и адаптируют свои характеристики в процессе взаимодействия с другими агентами [1-3]. Однако если проводить параллели между сообществом таких интеллектуальных агентов и биологическими системами, то видна необходимость улучшения каждым агентом своего функционирования в процессе существования популяции, накопления этой популяцией жизненного опыта и как следствие увеличение объема знаний каждым агентом. Здесь уже каждый агент (или группа агентов), как в централизованной системе управления, решает полностью некоторую задачу, но делает это параллельно с другими агентами (или группами). Поэтому популяция таких агентов в своей эволюции получает сразу столько решений стоящих перед ними (в общем случае различных) задач, сколько имеется в ней интеллектуальных агентов (или групп агентов). Это принципиальное отличие моделей искусственной жизни (в широком смысле этого понятия) от других эволюционных алгоритмов и МАС.

Рассматривается одна из возможных реализаций искусственной жизни, как эволюционирующей популяции интеллектуальных агентов. При этом цель разработки заключается в повышении качества (или сокращения времени вычислений) при решении множества подобных задач. То есть речь идет о параллельном решении некоторого множества задач множеством агентов, когда каждый из них находит решение и, как следствие, мы имеем в результате множество требуемых решений.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (гранты 02-01-00784, 03-07-90012).

1 Многоагентная система

При представлении сложной системы в виде МАС используется некоторое множество когнитивных агентов $\{A_1, A_2, \dots, A_k\}$. При этом для оптимального управления некоторой сложной системой с помощью данной МАС необходимым является оптимизация поведения каждого агента в МАС. Каждый i -ый агент характеризуется набором параметров $\{a_{i1}, a_{i2}, \dots, a_{in}\}$, которые и определяют его поведение в системе. Рассмотрим систему однотипных по структуре когнитивных агентов с одинаковым набором параметров:

$$\text{МАС} = \{A_1(a_{11}, \dots, a_{1n}), A_2(a_{21}, \dots, a_{2n}), \dots, A_k(a_{k1}, a_{k2}, \dots, a_{kn})\}.$$

В качестве примера такой системы можно привести производственный участок, состоящий из нескольких станков, транспортного робота, погрузочно-разгрузочного робота и конвейера. Здесь агент A_i ; моделирует i -ый станок, а его параметры $\{a_{i1}, a_{i2}, \dots, a_{in}\}$ определяют приоритеты стенка на вызов транспорта, на вызов погрузочно-разгрузочного робота, на погрузку деталей в тележку, на разгрузку деталей с конвейера и т.п.

Для решения задачи в МАС необходимо оптимизировать поведение каждого агента в системе. Это значит, что необходимо найти такую комбинацию параметров агента, при которой его поведение в системе будет оптимальным:

$$\text{МАС}^{\text{opt}} = \{A_1^{\text{opt}}, A_2^{\text{opt}}, \dots, A_k^{\text{opt}}\},$$

где

$$A_1^{\text{opt}} : A_1(a_{11}, a_{12}, \dots, a_{1n}) \rightarrow \text{opt};$$

$$A_2^{\text{opt}} : A_2(a_{21}, a_{22}, \dots, a_{2n}) \rightarrow \text{opt};$$

...

$$A_k^{\text{opt}} : A_k(a_{k1}, a_{k2}, \dots, a_{kn}) \rightarrow \text{opt}.$$

В основу модели положена эволюционная стратегия, включающая три механизма: воспроизведение, скрещивание и мутация. В некотором пространстве поиска создается популяция особей-агентов (их также можно назвать реактивными агентами), в которой один реактивный агент связан с одним когнитивным агентом МАС. У каждого реактивного агента имеется своя (отличная от других) целевая функция или функция пригодности, которую он оптимизирует в процессе своего развития в популяции. В формулу для расчета функции пригодности реактивного агента входят параметры связанного с ним когнитивного агента или параметры, зависящие от параметров когнитивного агента. Данная формула составляется таким образом, что при достижении оптимума функции пригодности реактивного агента, параметры когнитивного агента становятся оптимальными. Необходимо отметить, что все значения, которые может принимать функция пригодности каждого реактивного агента, лежат в пространстве поиска всей популяции реактивных агентов. Коллективное поведение реактивных агентов, их взаимодействие друг с другом посредством законов эволюции (скрещивание и мутация) и параллельное развитие (воспроизведение популяции) позволяют оптимизировать их функции пригодности и как следствие оптимизировать поведение когнитивных агентов исходной МАС.

Эволюционная стратегия реализуется как показано на рисунке 1. Блоки генерации первого поколения, скрещивания (используется одноточечный кроссинговер) и мутации реализуют операторы простого поискового генетического алгоритма [1]. Поэтому далее рассмотрим реализацию лишь блока воспроизведения популяции.

2 Блок воспроизведения популяции

Реализует воспроизведение особей и формирование новой популяции. Для каждой особи в последующую популяцию отбирается один лучший потомок (с наибольшей ФП), тем самым, продолжая «эволюционную линию» данной особи. Особь родитель при этом погибает. Таким образом, число особей в последующей популяции равно числу особей в предыдущей популяции и неизменно при переходе популяции от поколения к поколению. Воспроизведение осуществляется следующим образом:

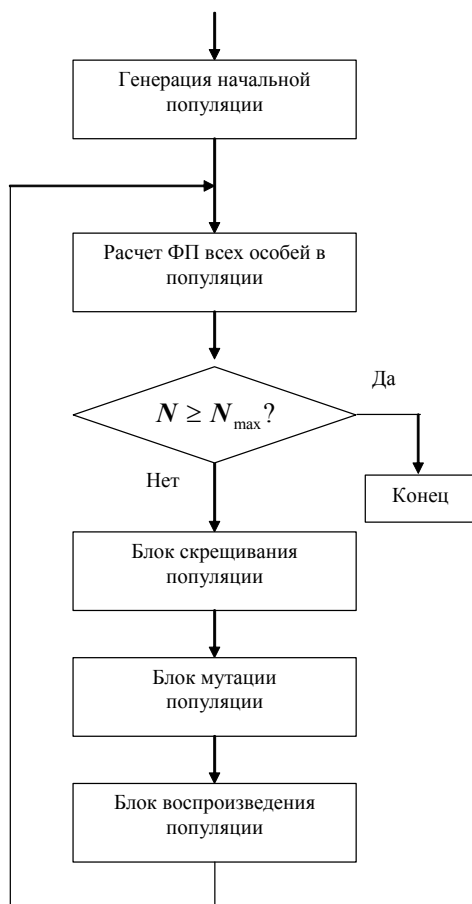


Рисунок 1 – Схема алгоритма эволюции.

Создается новая популяция особей, которая будет состоять из потомков особей текущей популяции. Причем в новую популяцию отбирается по одному потомку от каждой особи текущей популяции. Номер поколения N для новой популяции увеличивается на единицу по сравнению с номером поколения текущей популяции.

- Выбирается очередная особь в текущей популяции с номером i . Параметр i представляет собой значения счетчика изменяющегося в пределах от 1 до числа особей в популяции. Таким образом, перебираются по очереди все особи в популяции.
- Если Особь i имеет три потомка то выбирается лучший (с максимальной ФП) потомок. Если же особь имеет одного потомка, то он и является лучшим потомком i -ой особи.
- Потомок i добавляется в новую популяцию под номером i , тем самым, продолжая «эволюционную линию» i -ой особи.
- Проверяется условие окончания блока воспроизведения популяции. Если новая популяция полностью сформирована, то старая популяция удаляется. При этом «погибают» старые особи.

3 Модели искусственной жизни

Покажем работу модели «Искусственной жизни» на тестовых задачах.

Первая из задач заключается в следующем:

Имеется дискретное (с шагом 2-8) двумерное пространство с изменением координат от 0 до 1. В пространстве имеется 25 точек-оптимумов (для каждой особи-агента) равноудаленных друг от друга. В данном пространстве «живут» 25 особей-агентов и каждая особь-агент имеет свой оптимум, к которому стремится в результате развития популяции. Визуально пространство поиска показано на рисунке 2.

Для сравнения, эта же задача решается также методом случайного перебора, в котором: количество генераций случайных особей-агентов на каждом поколении равно числу агентов потомков в модели "Искусственная жизнь". Сделано это для того, чтобы уравнивать вычислительные ресурсы, затрачиваемые в каждой из моделей, что необходимо для сравнения результатов.

Сравнительные характеристики развития популяции для обеих моделей на 5, 20, 50 и 100 поколениях приведены на графиках. Здесь сравниваются максимальные значения ФП (рисунок 3) и средние значения ФП (рисунок 4).

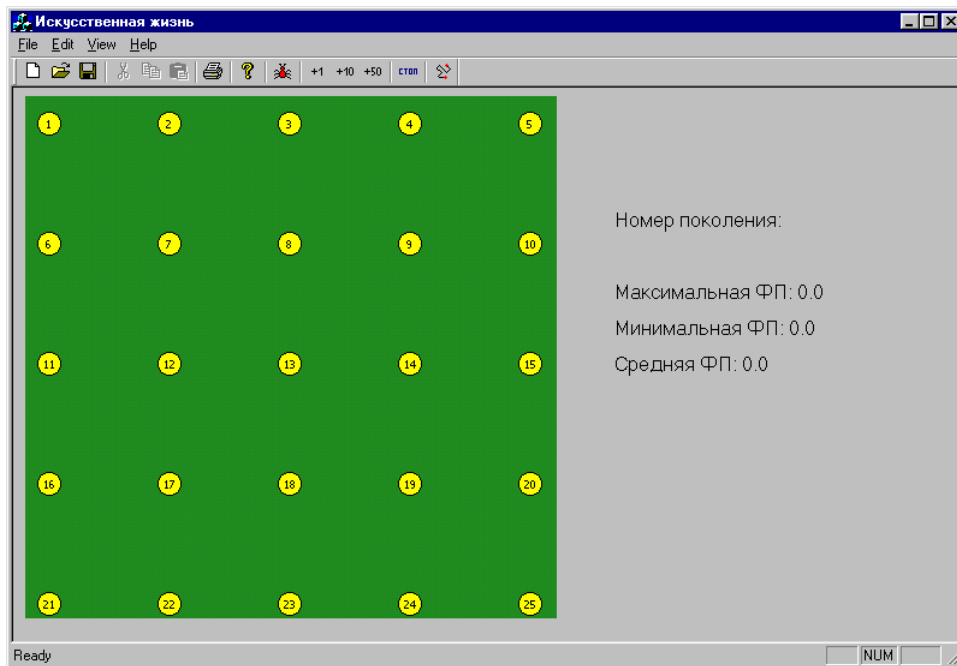


Рисунок 2 – Пространство поиска.

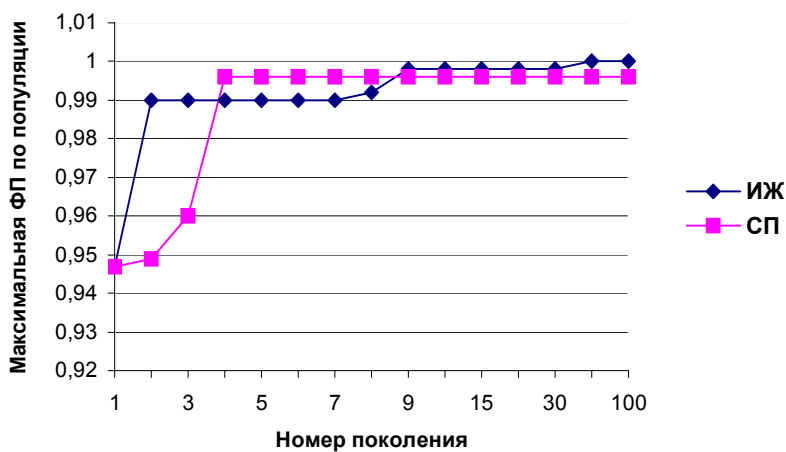


Рисунок 3 – Изменение максимальных значений ФП.

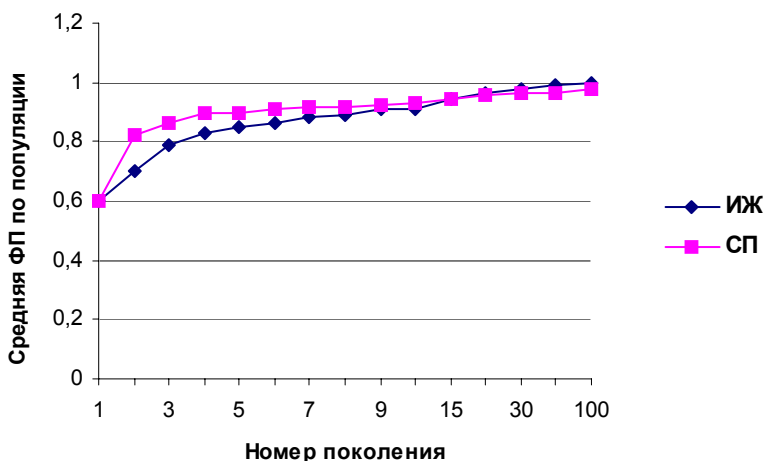


Рисунок 4 – Изменение средних значений ФП.

Из графиков видно, что модель Искусственной жизни более медленно выходит на установившееся функционирование, но при этом обеспечивает лучшее решение, чем случайный поиск. Следовательно, можно говорить о перспективности развития данного направления для решения однотипных задач популяцией интеллектуальных агентов.

Во второй задаче агенты в популяции принадлежат к одной из четырех групп (отличаются цветом). Задача агентов каждой группы – найти на площади неправильного четырехугольника точки, равноудаленные от вершин, имеющих цвет, отличный от цвета агентов группы. То есть каждая группа решает свою собственную задачу, но при этом обменивается генетическим материалом с другими группами. На рисунке 5 приведено расположение 20 агентов (по пять агентов различных цветов) в первом поколении популяции. На рисунках 6 и 7, соответственно приведены 11 и 41 поколения в эволюции популяции.

На рисунках приведены также графики изменения функции пригодности (максимальное значение, минимальное и среднее) для каждой из групп агентов. Из результатов видно, что агенты достаточно быстро находят решения близкие к оптимальным.

Заключение

Приведенные результаты позволяют сделать вывод о перспективности использования моделей искусственной жизни для решения оптимизационных задач.

Модели искусственной жизни имеют следующие особенности:

- 1) Особи (агенты) в популяции независимы – каждая из них имеет свой оптимум и решает свою оптимизационную задачу.
- 2) В процессе эволюции нас интересуют все особи популяции, в отличие от случаев, когда в итоге ищется лишь одна наилучшая особь, которая и есть решение единственной задачи.
- 3) Оптимум для особи (группы особей) может меняться в процессе эволюции. Динамическое изменение оптимумов позволяет моделировать воздействие внешней среды на эволюцию популяции.

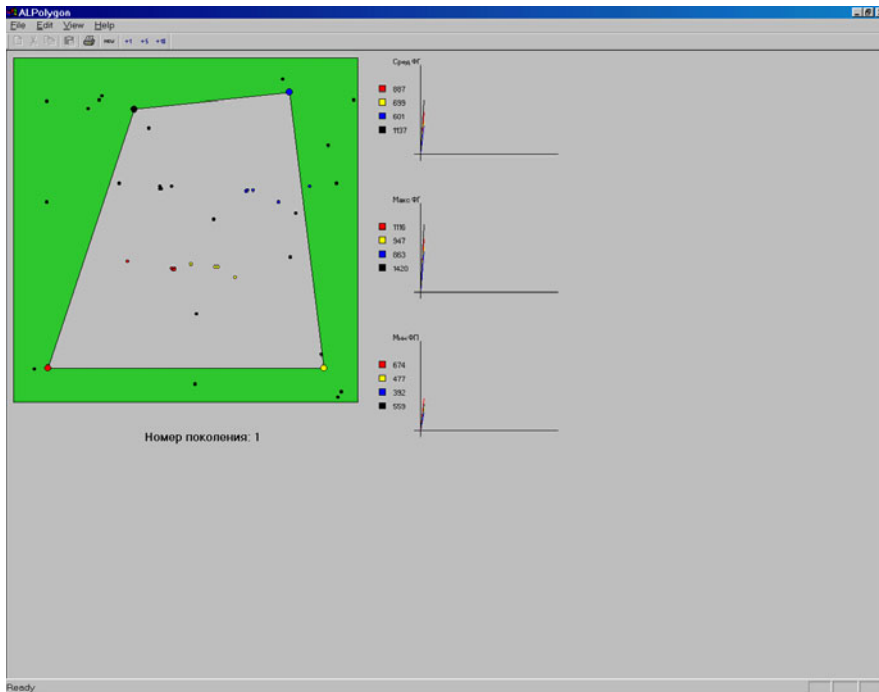


Рисунок 5 – Первое поколение популяции.

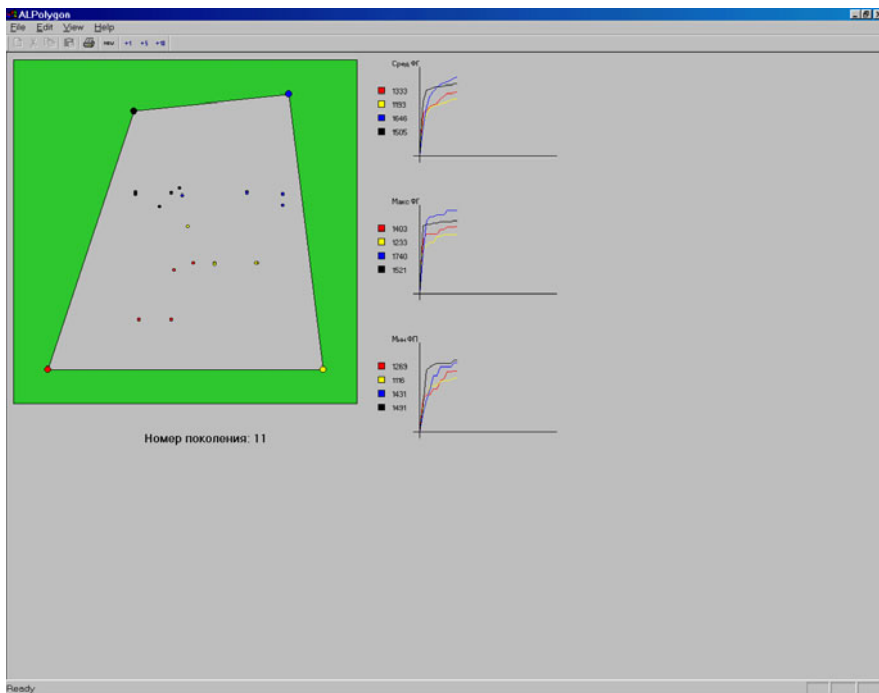


Рисунок 6 – Одиннадцатое поколение популяции.

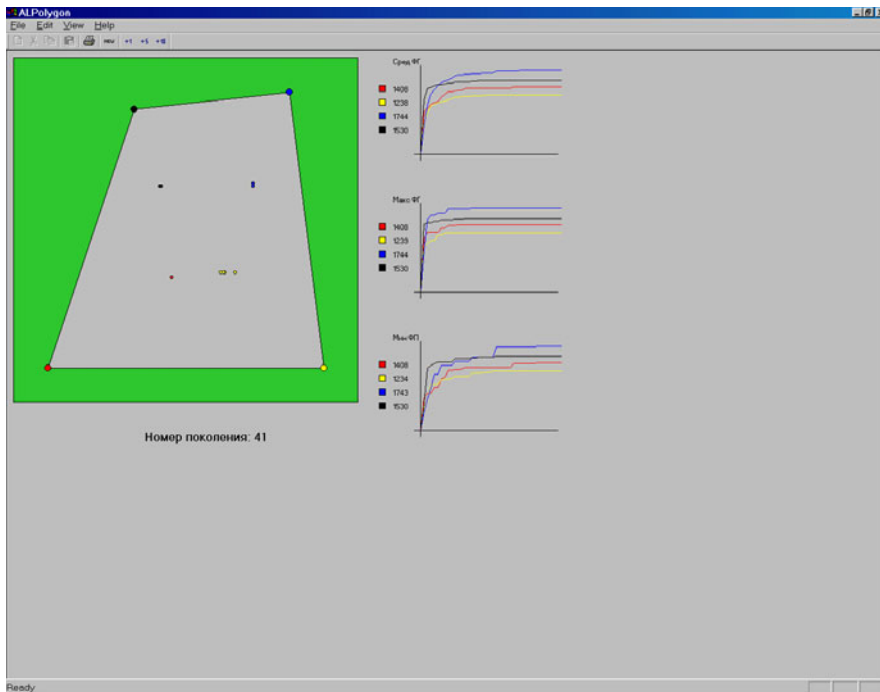


Рисунок 7 – Сорак первое поколение популяции.

Список литературы

- [1] Емельянов В.В., Курейчик В.М., Курейчик В.В. Теория и практика эволюционного моделирования. – М.: ФИЗМАТЛИТ, 2003.
- [2] Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002.
- [3] Редько В.Г. Эволюционная кибернетика. – М.: Наука, 2001.

МОДЕЛИ В СИСТЕМАХ АДАПТИВНОГО КОНТРОЛЯ ЗНАНИЙ

В.В. Емельянов¹, Ю.О. Шепитько²

¹МГТУ им. Н.Э.Баумана
107005, Москва 2-я Бауманская, 5
evv@rk9.bmstu.ru

тел: 263-66-39, факс (095) 263-61-70

²Таганрогский государственный радиотехнический университет
347928, Таганрог, пер. Некрасовский, 44, ТРТУ,

jush@mail33.com

тел: (8634) 31-14-87

Ключевые слова: *тест, уровень сложности, уровень знаний, модель, имитационное моделирование, обучение*

Abstract

Now competition between the manufacturers of educational services considerably grows. The organization of engineering education undergoes significant changes, with the purpose of the adaptation of high schools to new conditions of the market. The change of the purposes of education is necessary, in which on the foreground the task of development of the person with the help of individual training leaves. One of ways of achievement of such purposes is the organization of remote education, creation of virtual faculties, laboratories, universities. Among problems, arising at it, there is an important problem connected to an estimation of knowledge of testing, trained on the basis of computer systems. The importance of correct organization of testing of knowledge trained with the purpose of their estimation considerably grows in remote education, when the teacher and tested are far apart. The now in use circuits of testing have many lacks. Among them it is possible to note absence of a feedback between results of testing and developers of the tests, static character of models (for example, Rasch's model, Birnbaum's model, etc.), on which the processing of results, absence of the account of change of a level of knowledge in time is carried out and during testing, impossibility of forecasting of results and others. Use of simulation as a means for planning process of testing and account of dynamics in its realization removing a part from the specified problems, is considered below.

Введение

В настоящее время значительно возрастает конкуренция между производителями образовательных услуг. Организация инженерного образования претерпевает значительные изменения. Необходимо изменение целей образования, в которых на первый план выходит задача развития личности с помощью индивидуального обучения [1]. Одним из путей достижения таких целей является организация дистанционного образования, создание виртуальных кафедр, лабораторий, университетов [2]. Среди возникающих при этом проблем возникает важная проблема, связанная с оценкой знаний обучаемых на основе компьютерных систем тестирования. Значимость правильной организации контроля знаний обучаемых с целью их оценки значительно возрастает в дистанционном образовании, когда преподаватель и тестируемые находятся далеко друг от друга. Используемые в настоящее время схемы тестирования имеют достаточно много недостатков. Среди них можно отметить отсутствие обратной связи между результатами тестирования и разработчиками тестов, статичность моделей (например, модели Раша, Бирнбаума и др.), на которых осуществляется обработка результатов, отсутствие учета изменения уровня знаний во времени и в процессе тестирования, невозможность прогнозиро-

вания результатов и ряд других. Использование имитационного моделирования снимает часть из указанных проблем.

Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (гранты 02-01-00784, 02-01-81015).

1 Адаптивное тестирование

Среди важных интерактивных функций программного обеспечения дистанционного образования и виртуальных кафедр особое место занимают средства обеспечивающие проведение преподавателем тестирования знаний без непосредственного взаимодействия с обучаемыми [3-5]. Интеллектуальная обучающая система должна обладать возможностью глубокого анализа ответов обучаемого при тестировании знаний и возможность подготовки тестов. С этой целью в ее состав включается подсистема тестирования знаний.

Адаптивное тестирование - это широкий класс методик тестирования, предусматривающих изменение последовательности предъявления заданий в самом процессе тестирования с учетом ответов испытуемого на уже предъявленные задания. В узком смысле к АТ обычно относят особые алгоритмы предъявления заданий, построенные для пунктов теста, предварительно отобранных с помощью соответствующих моделей и методов анализа пунктов и других процедур, основанных на психометрической теории задание-ответ (Item Response Theory). При таком обучении в процессе прохождения теста (или набора тестов) строится модель обучаемого, которая используется для генерации или выбора последующих заданий тестирования в зависимости от уровня обучаемого. В комплексных системах полученная модель также может использоваться в процессе обучения.

В качестве средства для планирования процесса тестирования и учета фактора времени в его проведении предлагается использовать имитационное моделирование. Т.е. подсистема тестирования должна содержать модель или ряд моделей, имитирующих процесс тестирования. Их использование позволит учесть ряд факторов, таких как время решения задания и время, отведенное на тест, утомляемость тестируемого, изменение уровня его знаний и т.п. При этом появляется возможность оптимизации тестового задания – выбора количества заданий и времени, отведенного на их выполнение, выбора стратегии выполнения теста, и других задач связанных с планированием процесса тестирования. В качестве варьируемых факторов могут выступать: диапазон уровней знаний тестируемых, диапазон трудностей заданий, степень соответствия этих диапазонов, время тестирования, число вопросов (заданий) в тесте. В качестве входных данных должны использоваться задания, отличающиеся числом и уровнем трудности задач, множество тестируемых имеющих различный уровень подготовленности, данные об изменении степени утомляемости тестируемого в процессе его выполнения и др. На выходе получаем информацию о результатах тестирования.

Очевидно, что в процессе тестирования у тестируемого происходит последовательное повышение уровня понимания. Поэтому адаптивная подсистема тестирования знаний, должна динамически отслеживать изменение состояния тестируемого и активно изменять сложность теста в зависимости от текущего состояния тестируемого.

2 Математические модели теории тестов

Рассмотрим ряд моделей теории тестов с целью выявления характеристик тестов и тестируемых, которые должны быть учтены и в имитационных моделях.

Под современной теорией тестов понимается теория, предназначенная для оценки латентных параметров N тестируемых θ_i , ($i = 1, 2, \dots, N$), и параметров заданий теста β_j , ($j = 1, 2, \dots, M$) посредством применения математико-статистических моделей измерения. Она позволяет получить устойчивые и объективные оценки параметров, характеризующих

уровень подготовки тестируемых и трудности заданий, а также измерять эти параметры в одной и той же шкале, имеющей свойства интервальной. Это должно обеспечить возможность предсказания вероятности правильного выполнения теста любым тестируемым в выборке до предъявления теста группе тестируемых. По их ответам на задания теста необходимо оценить значения латентных параметров θ_i и β_j [6].

Г. Раш для установления соотношения между латентными параметрами θ_i и β_j ввел в рассмотрение их разность. Тогда величина $|\theta_i - \beta_j|$ – это расстояние от уровня подготовки тестируемого от задания с трудностью β_j . Однопараметрическая модель Раша позволяет определить вероятностные функции распределения латентных параметров:

$$P_j(\theta) = \frac{1}{1 + e^{-1,7(\theta - \beta_j)}}, \quad P_i(\beta) = \frac{1}{1 + e^{-1,7(\theta_i - \beta)}}$$

где θ и β – независимые переменные для первой и второй логистической функции.

А. Бернбаум ввел в модель параметр a_j для характеристики дифференцирующей способности задания при различных изменениях значений θ и параметр a_i как меру структурированности знаний тестируемого. Двухпараметрическая модель Бернбаума имеет вид:

$$P_j(\theta) = \frac{1}{1 + e^{-1,7a_j(\theta - \beta_j)}}, \quad P_i(\beta) = \frac{1}{1 + e^{-1,7a_i(\theta_i - \beta)}}$$

Более сложной является трехпараметрическая модель Бернбаума, где третий параметр c_j характеризует вероятность правильного ответа на задание j в том случае, если этот ответ угадан:

$$P_j\{x_{ij} = 1 | \beta_j\} = c_j + (1 - c_j) \frac{1}{1 + e^{-1,7a_j(\theta - \beta_j)}},$$

где

$$x_{ij} = \begin{cases} 1, & \text{если ответ } i\text{-го тестируемого на } j\text{-ое задание верный;} \\ 0, & \text{если ответ } i\text{-го тестируемого на } j\text{-ое задание неверный.} \end{cases}$$

Приведенные модели по своей сути являются статическими, то есть их параметры постоянны во времени. Очевидно, что при дистанционном тестировании, когда процесс тестирования является достаточно протяженным, большинство параметров человека, проходящего тестирование, изменяется. Поэтому необходимо иметь возможность по текущим результатам теста проводить коррекцию его заданий, то есть осуществлять обратную связь между преподавателем (интеллектуальной обучающей системой) и тестируемыми.

Кроме указанного недостатка, приведенные модели не учитываются ряд параметров тестируемого, которые влияют на результат испытаний. Среди последних можно указать: базовую информацию о тестируемом, характеризующую его индивидуальность (имя, пол, характер, возраст, способности); социальный фактор (городской, деревенский, из богатой семьи, из бедной семьи, из многодетной семьи); стороны жизни (родители с высшим образованием, родители рабочие, родители алкоголики); вид досуга (спорт, компьютеры, гулянки, учебники); способ подготовки (заочные школы, дополнительные курсы, самоподготовка, репетитор); здоровье (крепкое, среднее, слабое) и многие другие.

3 Имитационная модель процесса тестирования

Рассмотрим одну из имитационных моделей, которая может быть использована при тестировании. Модель разрабатывается в среде РДО, которая является универсальной средой мо-

делирования, основанной на формализованном описании знаний о предметной области [7]. Эти знания описывают ресурсы (под которыми могут пониматься любые элементы моделируемых системы или процесса) и те правила, по которым эти ресурсы взаимодействуют между собой во времени и пространстве.

Среди основных типов (классов) ресурсов модели будем рассматривать задания, учеников и этапы тестирования. Тест, предъявляемый ученикам, включает ряд заданий.

```
$Resource_type Задания : permanent
$Parameters
балл          : real = 0.0 {проверка сложности-легкости задания}
состояние     : (не_выполнено, выполнено, выполняется, редактируется) = не_выполнено
сложность     : integer = 0
время_выполнения : real = 0.0
выполнили_задание : integer = 0
$End
```

Параметр *балл* содержит информацию о том, сколько учеников успешно справилось с данным заданием. Параметр *состояние* идентифицирует, что в данный момент происходит с заданием; изначально все задания не выполнены. Параметр *сложность* показывает текущую сложность задания. Параметр *время_выполнения* содержит значение времени, которое тратится на решение данного задания. Параметр *выполнили_задание* показывает сколько учеников уже решали это задание.

```
$Resource_type Ученики : permanent
$Parameters
состояние : (новичок,обучился, решает_тест, свободен, выполнил_весь_тест, обучается,
            выполнил1, выполнил2, выполнил3, выполнил4, выполнил5,
            выполнил6, выполнил7, выполнил8, выполнил9, выполнил10,
            выполнил11, выполнил12, выполнил13, выполнил14, выполнил15) = новичок
социальный_фактор : (городской, деревенский, из_богатой_семьи, из_бедной_семьи, из_многодетной_семьи)
стороны_жизни    : (родители_с_ВО, родители_рабочие, родители_алкоголики)
вид_досуга       : (спорт, компьютеры, гулянки, пиво, учебники)
способ_подготовки : (заочные_школы, дополнительные_курсы, самоподготовка, репетитор)
здоровье         : (крепкое, среднее, слабое)
суммарный_балл   : integer = 0
вероятность_решения : real = 0.0
получил          : integer = 0
$End
```

Параметр *состояние* показывает на какой стадии находится ученик, изначально все ученики помечаются как новички, впоследствии этот параметр нигде не используется, он необходим чтобы определить изначальный уровень подготовки тестируемых. Параметры *социальный_фактор*, *стороны_жизни*, *вид_досуга*, *способ_подготовки*, *здоровье* характеризуют каждого отдельного ученика; исходя из этих данных производится подсчет первоначальных знаний тестируемых. Параметр *суммарный_балл* показывает количество баллов, набранных учеником за время выполнения теста. Параметр *вероятность_решения* показывает изначальный уровень знаний учеников, а в последствии после обучения и их текущие знания. Параметр *получил* используется во время моделирования и показывает что получил (0 или 1) ученик за конкретное задание.

```
$Resource_type Этап : permanent
$Parameters
Количество_выполненных_тестов : integer = 0
средний_балл_за_тесты          : real = 0.0
заданий_отредактировано_легких : integer = 0
заданий_отредактировано_сложных : integer = 0
число_учеников_выше_среднего    : integer = 0
число_учеников_ниже_среднего    : integer = 0
$End
```

Одно из правил модели, а именно правило выполнения задания с номером 15, имеет вид приведенный ниже. На выполнение задания отводится 30 минут времени, а полученный балл зависит сложности задания и параметров ученика:

```

$Pattern выполнение_задания15 : operation
$Relevant_resources
задание : Задание15 Keep Keep
уч : Ученики Keep Keep
цифры : данные NoChange Keep
расчет : показатели NoChange Keep
$Time = 30.0
$Body
задание
Choice from задание.состояние = не_выполнено
first
Convert_begin
состояние set выполняется
Convert_end
состояние set не_выполнено
выполнили_задание set задание.выполнили_задание + 1
уч
Choice from уч.состояние = выполнил14
first
Convert_begin
состояние set решает_тест
Convert_end
состояние set выполнил15
получил set решение_задания(уч.вероятность_решения*(15-задание.сложность)*10,решение(0.0,100.0))
суммарный_балл set уч.суммарный_балл + уч.получил
цифры
Choice from цифры.счетчик14 = 15 and цифры.счетчик15 <= 15
first
Convert_end
счетчик15 set цифры.счетчик15 + 1
расчет
Choice from расчет.сумма_за15 <= 15 and расчет.стоп = 1
first
Convert_end
сумма_за15 set расчет.сумма_за15 + уч.получил
$End

```

В приведенном правиле учет параметров ученика реализован с помощью функции *вероятность_решения*, которая имеет вид:

```

$Function вероятность_решения_заданий1 : real
$Type = algorithmic
$Parameters
фактор1 : such_as Ученики.социальный_фактор
$Body
Calculate_if фактор1 = городской вероятность_решения_заданий1 = 0.8
Calculate_if фактор1 = деревенский вероятность_решения_заданий1 = 0.4
Calculate_if фактор1 = из_богатой_семьи вероятность_решения_заданий1 = 0.75
Calculate_if фактор1 = из_бедной_семьи вероятность_решения_заданий1 = 0.65
Calculate_if фактор1 = из_многолетней_семьи вероятность_решения_заданий1 = 0.7
$End

```

Приведенные элементы дают некоторое представление о том, как в имитационной модели описывается процесс тестирования.

4 Результаты моделирования

В результате ряда прогонов модели с разными параметрами и различными базами генераторов случайных чисел, были собраны данные, далее обработанные при помощи программы MS Excel.

На рисунке 1 показано как изменяется средний балл тестируемых во времени по этапам тестирования. Сплошная линия соответствует среднему значению, а штриховая это отклонение от среднего. Из графика видно, что с увеличением числа этапов средний балл стремится к определенному значению.

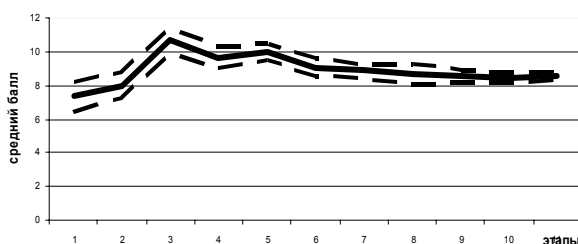


Рисунок 1 – Средний балл за тест.

На следующем рисунке 2 показана зависимость знаний тестируемых (средних знаний по всем тестируемым) в зависимости от этапов тестирования. Под знаниями в модели понимается вероятность выполнения данного теста. Сплошной линией показано среднее значение вероятности решения заданий, а штриховой линией отклонение от среднего. Здесь можно заметить, что отклонение от среднего имеет большой разброс, это обусловлено, тем в каждой точке подчитывается среднее значение 15 учеников, а по отдельно взятым ученикам диаграмма может иметь совершенно другой вид и направленность. Но в целом можно заметить, что с увеличением числа решенных тестов, среднее значение вероятности решения растет, это говорит о том, что с каждым этапом задания все больше соответствуют уровню знаний тестируемых.

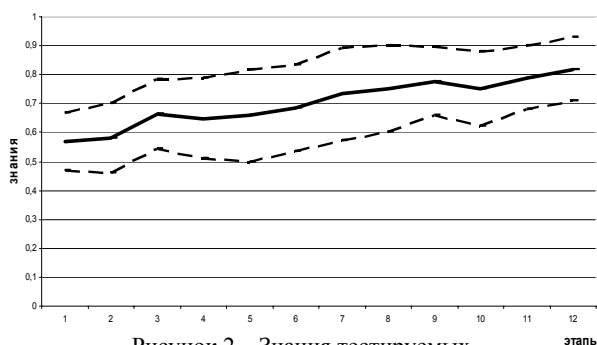


Рисунок 2 – Знания тестируемых.

На рисунке 3 приведен график числа тестируемых со средними знаниями по итогам текущего теста, а также число тестируемых, со знаниями выше и ниже среднего.

Из графика можно заметить, что с каждым следующим этапом тестирования число тестируемых со средними знаниями стремится к определенному уровню, в то время как количество учеников выше и ниже среднего соответственно увеличивается и уменьшается.

Заключение

В данной работе представлены некоторые результаты и идеи разработки имитационной модели для использования в адаптивной тестирующей подсистеме, являющейся частью интеллектуальной системы обучения. Особенностью является интеграция известных подходов к тестированию и имитационного моделирования. Это должно позволить проводить адаптивное тестирование с учетом динамики характеристик тестируемого, планировать разработку эффективных тестов и прогнозировать результаты тестирования. Приведенные результаты, позволя-

ют сделать вывод о том, что имитационная модель обеспечивает адекватное описание процесса с требуемым для адаптации уровнем детализации.

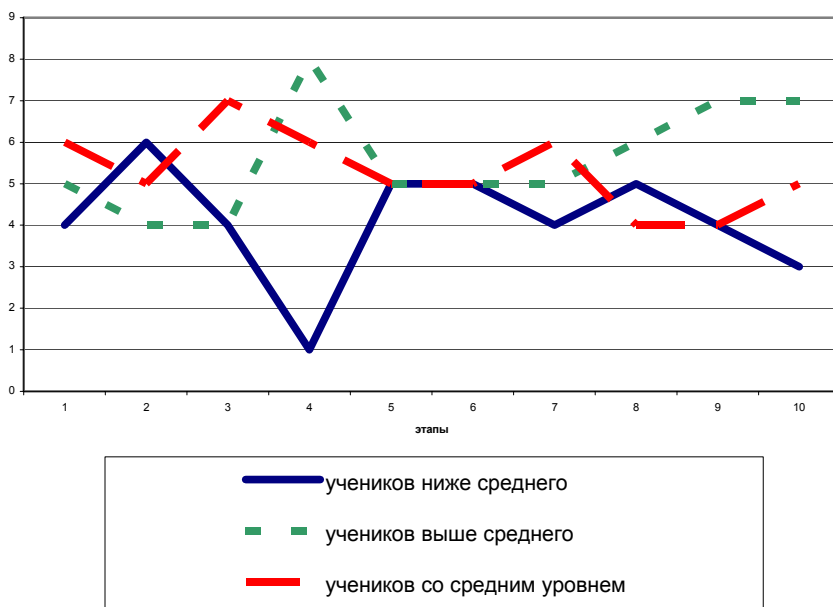


Рисунок 3 – Распределение тестируемых.

Список литературы

- [1] Равен Д. Педагогическое тестирование: Проблемы, заблуждения, перспективы / Пер. с англ. Изд. 2-е, испр. – М.: Когнито-центр, 2001.
- [2] Интеллектуальные обучающие системы и виртуальные учебные организации: Монография / В.В. Голенков, В.Б. Тарасов, О.Е. Елисеева, В.В. Емельянов и др.: Под ред. В.В. Голенкова и В.Б. Тарасова. – Мн.: БГУИР, 2001.
- [3] Емельянов В.В., Тарасов В.Б. Виртуальная кафедра в техническом университете // Дистанционное образование. 2000. №6. С.39-45.
- [4] Barron D. Distance Education: the Virtual Classroom Updated // School Library Media Annual. 1994. Vol. 12.
- [5] Hillz S.R. Correlates of Learning in Virtual Classroom // Int. Journal of Man-Machine Studies. 1993. Vol. 39. P. 71-98.
- [6] Челышкина М.Б. Теория и практика конструирования педагогических тестов: Учебное пособие. – М.: Логос, 2002.
- [7] Емельянов В.В., Ясиновский С.И. Введение в интеллектуальное имитационное моделирование сложных дискретных систем и процессов. Язык РДО. – М.: АНВИК, 1998.

ДИНАМИЧЕСКАЯ МОДЕЛЬ САМООРГАНИЗУЮЩЕГОСЯ НОМА

В.К. Захаров¹, А.Д. Яшин²

¹Московский государственный университет им. М.В. Ломоносова
119992, Москва, Ленинские Горы
valeri@zaharov.mccme.ru

тел: . (8095) 113-87-04, факс (8095) 939-44-41

²Московский городской психолого-педагогический университет
127051, Москва, ул. Сретенка, 29
yashin@mgppu.ru

тел: (8095) 345-46-08, факс: (8095) 9234896

Ключевые слова: *жизнеорганизующий цикл, балансовые уравнения, импульсные уравнения*

Abstract

In this paper a dynamic model of a *nomós* (i.e, a linked vitally active community) is introduced. The structure of a nomos is described in terms of the three basic archetypes of vital activity: *substantiating*, *directing*, and *ensuring*. The vital activity of a nomos is presented in the form of a vitally organizing cycle of internal streams of results of vital activity linking units and main parts of the nomos with each other and with the *natural environment* and the *external organized environment*. Some natural conditions for resources of units and streams between them are given and some balance and impulse equations for them are proposed.

Введение

Человечество в процессе жизнедеятельности творит собственную новую среду жизнедеятельности: и материальную, и духовную – и собственных новых субъектов жизнедеятельности. При этом предыдущая жизнедеятельность наследуется человечеством в форме *архетипов жизнедеятельности*. Эти архетипы, соответствующим образом воздействуя на жизнедеятельность в настоящем, направляют также дорогу в будущее. Таким образом, **исключается абсолютная произвольность** новой жизнедеятельности.

Чем дальше в прошлое уходит складывание и проявление какого-либо архетипа жизнедеятельности, тем большее влияние он оказывает на настоящее и будущее. Однако с развитием во времени человечество оставляет за собой все больший исторический шлейф, который приводит к оформлению все большего числа всё более молодых архетипов жизнедеятельности. Поэтому прибавление этих более молодых архетипов влияет на проявление более старых, что приводит к изменению этого проявления. Этим исключается **абсолютная предопределенность** новой жизнедеятельности.

Можно ввести понятие *комплекса архетипов* жизнедеятельности, близкое к понятию генного комплекса в биологии. Комплекс архетипов, исключая и абсолютную произвольность, и абсолютную предопределенность, задаёт *тенденции жизнедеятельности*, от которых жизнедеятельность может уклоняться во времени лишь под воздействием новых общественных сил пропорционально величине этого воздействия.

Таким образом, вычленение исследователем архетипов жизнедеятельности способствует не только пониманию прошлого и настоящего, но и дает некоторую надежду на предвидение ближайшего будущего.

Работа посвящена выделению и математическому описанию *архетипа организационного жизнеустройства человечества*. Согласно этому архетипу жизнедеятельность человечества в основном осуществляется в динамически связанных жизнедеятельных общностях, называемых в этой работе *номами*.

1 Ноновая система

Рассмотрим конечное множество U единиц нома и две другие единицы a_1 и a_2 , называемые *природной средой* и *внешней организованной средой*. Введем разбиение $U = E \cup B \cup P \cup H$ множества U на следующие множества:

- E - множество *содержательных единиц*,
- B - множество *распорядителей содержательных единиц*,
- P - *властных распорядительных единиц*,
- $H \equiv \bigcup(H_i \mid i \in I)$ - множество *всех обеспечительных единиц* и
- H_i - i -е множество *обеспечительных единиц*.

Множества E , B , P и H назовем *основными частями нома*, множество $F \equiv E \cup B$ - *основанием нома*, а множество $P \cup H$ - *организационной надстройкой нома*.

Зафиксируем непустое подмножество $\overline{{}^2P} \subset {}^2P \equiv \{\{p, q\} \mid p, q \in P \wedge p \neq q\}$ такое, что пара $(P, \overline{{}^2P})$ образует граф. На P имеется строгий частичный порядок $<$.

Пусть $MaxP$ [$MinP$] – множество всех максимальных [минимальных] единиц P и $\overline{\overline{{}^2P}} \equiv \overline{{}^2P} \setminus \{\{p, q\} \in \overline{{}^2P} \mid p, q \in MaxP\}$ - *усеченное* множество ребер графа. Граф и порядок связаны условиями:

- 1) подграф $(P, \overline{{}^2P})$ является деревом;
- 2) $\forall p, q \in P(\{p, q\} \in \overline{{}^2P} \Rightarrow p, q \in MaxP \vee p < q \vee q < p)$;
- 3) $\forall q \in MaxP \exists p \in P(\{p, q\} \in \overline{{}^2P} \wedge p < q)$;
- 4) $\forall q \in MaxP \forall p \in P(\{p, q\} \in \overline{{}^2P} \Rightarrow p \in MaxP \vee p < q)$;
- 5) $\forall p \in MinP \exists q \in P(\{p, q\} \in \overline{{}^2P})$;
- 6) $\forall p \in MinP \forall q \in P(\{p, q\} \in \overline{{}^2P} \Rightarrow p < q)$.

Систему $(P, \overline{{}^2P}, <)$, удовлетворяющую условиям 1)-6), называем *корневой системой властных распорядителей*. Аналогичным образом для $i \in I$ заданы *корневые системы обеспечителей* $(H_i, \overline{{}^2H_i}, <)$.

Имеется коллекция $(B_e \subset B \mid e \in E)$ такая, что $B = \bigcup(B_e \mid e \in E)$. Множество B_e называется *совокупностью распорядителей единицы $e \in E$* . Эта коллекция задает новые ребра $\{e, b\}$ для всех $e \in E$ и $b \in B_e$.

Имеется также коллекция $(\varphi_e \mid e \in E)$ такая, что $\varphi_e : B_e \rightarrow P$ и $MinP \subset \bigcup(rng \varphi_e \mid e \in E)$. Она задает новые ребра $\{b, \varphi_e(b)\}$ для всех $e \in E$ и $b \in B_e$.

Для P и H_i имеется соответствие $\zeta_i \subset P \times H_i$, для которого существует хотя бы один элемент $(p, h) \in \zeta_i$ такой, что $p \in \tilde{P}_i \equiv dom \zeta_i$, $h \in \tilde{H}_i \equiv rng \zeta_i$, $p \in MaxP$ и $h \in MaxH_i$. Оно задает новые ребра $\{p, h\}$ для всех $p \in P$ и $h \in H_i$ таких, что $(p, h) \in \zeta_i$.

Для множеств E и B заданы произвольные графы $(E, \overline{{}^2E})$ и $(B, \overline{{}^2B})$.

Все $u \in U$ вместе с a_1 и a_2 задают новые ребра $\{u, a_1\}$ и $\{u, a_2\}$.

В результате получается расширенный связный граф. Математическую систему $S \equiv (U; \sigma)$ с носителем U и структурой σ , определенной выше, назовем *номовой системой*.

2 Достояния единиц нома

Пусть $T \equiv [0, \infty[$ - временной интервал и $t \in T$. Каждой единице $u \in U$ приписаны три функции w_u^1, w_u^2, w_u^3 из T в \mathbf{R} , называемые, соответственно, *содержательным, распорядительным* и *обеспечительным* *достояниями* единицы u на момент времени $t \in T$. Далее со словами «содержательный», «распорядительный» и «обеспечительный» связываем цифры 1, 2 и 3, соответственно.

Достояния единиц обладают следующим свойством *монотонности*:

- 1) $w_b^m(t) < w_p^m(t)$ для любого t , любого $m = 1, 2, 3$ и любых $b \in B$ и $p \in P$, для которых существует $e \in E$ такое, что $p = \varphi_e(b)$;
- 2) $w_p^m(t) < w_q^m(t)$ для любого t , любого $m = 1, 2, 3$ и любых $p, q \in P$ таких, что $\{p, q\} \in \overline{{}^2P}$ и $p < q$;
- 3) $w_h^m(t) < w_k^m(t)$ для любого t , любого $i \in I$, любого $m = 1, 2, 3$ и любых $h, k \in H_i$ таких, что $\{h, k\} \in \overline{{}^2H_i}$ и $h < k$.

3 Потоки преобразований достояний единицы

Между содержательным $w_u^1(t)$, распорядительным $w_u^2(t)$ и обеспечительным $w_u^3(t)$ достояниями единицы $u \in U$ на момент времени t действуют *потоки преобразований* заданной интенсивности $f_u^{mn}(\tau)$ ($m, n = 1, 2, 3$) (напомним, что *интенсивностью потока* $A(\tau)$ называется его производная $\alpha(\tau) = \dot{A}(\tau)$ по времени τ):

- 1) для $e \in E$ заданы интенсивности $f_e^{11}(\tau)$, $f_e^{21}(\tau)$ и $f_e^{31}(\tau)$;
- 2) для $b \in B$ заданы интенсивности $f_b^{12}(\tau)$, $f_b^{22}(\tau)$ и $f_b^{32}(\tau)$;
- 3) для $p \in P$ заданы интенсивности $f_p^{12}(\tau)$, $f_p^{22}(\tau)$ и $f_p^{32}(\tau)$;
- 4) для $h \in H$ заданы интенсивности $f_h^{13}(\tau)$, $f_h^{23}(\tau)$ и $f_h^{33}(\tau)$.

4 Потоки результатов взаимодействия единиц

Между $u \in U$ и $v \in \{a_1, a_2\}$ по ребрам $\{u, v\}$ проходят m -потоки результатов внешнего взаимодействия единицы u с единицами a_1 и a_2 интенсивности $d_{uv}^m(\tau)$ и $d_{vu}^m(\tau)$, изменяющие m -достояние единицы u .

Между u и v из U по ребрам $\{u, v\}$ проходят потоки результатов внутреннего взаимодействия единиц u и v и заданы некоторые из интенсивностей $s_{uv}^{mn}(\tau)$ mn -потоков результатов взаимодействия, проходящих от единицы u к единице v по ребру $\{u, v\}$ и преобразующих m -достояние в n -достояние.

Точнее, предполагается, что

- 1) для $u \in U$ и a_1 заданы только $d_{a_1 u}^1(\tau)$; для $e \in E$ заданы только $d_{ea_1}^1(\tau)$; для $b \in B$ заданы только $d_{ba_1}^1(\tau)$; для $p \in P$ заданы только $d_{pa_1}^2(\tau)$ и для $h \in H$ заданы только $d_{ha_1}^3(\tau)$

(правило запрещения единицам $u \in U$ «сбрасывать» остальные виды своего достояния в природную среду);

- 2) для $u \in U$ и a_2 заданы все $d_{a_2 u}^m(\tau)$ и $d_{u a_2}^m(\tau)$ для $m = 1, 2, 3$;
- 3) для $e, f \in E$ таких, что $\{e, f\} \in \overline{{}^2 E}$, заданы все $s_{ef}^{mn}(\tau)$ и $s_{fe}^{nm}(\tau)$ для $m, n = 1, 2, 3$;
- 4) для $b, c \in B$ таких, что $\{b, c\} \in \overline{{}^2 B}$, заданы все $s_{bc}^{mn}(\tau)$ и $s_{cb}^{nm}(\tau)$ для $m, n = 1, 2, 3$;
- 5) для $e \in E$ и $b \in B_e$ заданы только $s_{eb}^{11}(\tau)$, $s_{be}^{22}(\tau)$ и $s_{be}^{33}(\tau)$;
- 6) для $e \in E$, $b \in B_e$ и $p \equiv \varphi_e(b) \in P$ заданы только $s_{bp}^{11}(\tau)$, $s_{pb}^{22}(\tau)$ и $s_{pb}^{33}(\tau)$;
- 7) для $p, q \in P$ таких, что $\{p, q\} \in \overline{{}^2 P}$ и $p < q$, заданы только $s_{pq}^{11}(\tau)$, $s_{qp}^{22}(\tau)$ и $s_{qp}^{33}(\tau)$;
- 8) для $p \in \tilde{P}_i$ и $h \in \tilde{H}_i$ такой, что $(p, h) \in \zeta_i$, заданы только $s_{ph}^{11}(\tau)$, $s_{ph}^{22}(\tau)$ и $s_{hp}^{33}(\tau)$;
- 9) для $h, k \in H_i$ таких, что $\{h, k\} \in \overline{{}^2 H_i}$ и $h < k$, заданы только $s_{kh}^{11}(\tau)$, $s_{kh}^{22}(\tau)$ и $s_{hk}^{33}(\tau)$.
- 10) для $p, q \in \text{Max}P$ таких, что $\{p, q\} \in \overline{{}^2 P}$, заданы все $s_{pq}^{mn}(\tau)$ и $s_{qp}^{nm}(\tau)$ для $m, n = 1, 2, 3$;
- 11) для $h, k \in \text{Max}H_i$ таких, что $\{h, k\} \in \overline{{}^2 H_i}$, заданы все $s_{hk}^{mn}(\tau)$ и $s_{kh}^{nm}(\tau)$ для $m, n = 1, 2, 3$.

5 Правило распределения всех потоков в единице

Если в $u \in U$ действует некоторый преобразовательный mn -поток $f_u^{mn}(\tau)$, то часть его остается в u и прибавляется к ее n -достоянию. Для этого постулируется, что заданы коэффициенты накопления $\kappa_u^{mn}(\tau) \leq 1$ ($m, n = 1, 2, 3$), которые для mn -потока $f_u^{mn}(\tau)$ задают накопительную долю $\kappa_u^{mn}(\tau)f_u^{mn}(\tau)$ единицы u от этого потока.

Если между $u \in U$ и $v \in \{a_1, a_2\}$ действует некоторый внешний m -поток $d_{vu}^m(\tau)$, то часть его остается в u и прибавляется к ее m -достоянию. Для этого постулируется, что заданы коэффициенты накопления $\gamma_{vu}^m(\tau) \leq 1$ ($m = 1, 2, 3$), которые для m -потока $d_{vu}^m(\tau)$ задают накопительную долю $\gamma_{vu}^m(\tau)d_{vu}^m(\tau)$ единицы u от этого потока.

Если из $u \in U$ в $v \in U$ действует некоторый внутренний mn -поток $s_{uv}^{mn}(\tau)$, то часть его остается в v и прибавляется к ее n -достоянию. Для этого постулируется, что заданы коэффициенты накопления $\lambda_{uv}^{mn}(\tau) \leq 1$ ($m, n = 1, 2, 3$), которые для mn -потока $s_{uv}^{mn}(\tau)$ задают накопительную долю $\lambda_{uv}^{mn}(\tau)s_{uv}^{mn}(\tau)$ единицы v от этого потока.

Все указанные коэффициенты являются первичными параметрами.

Правило распределения всех потоков в единице $v \in U$: если для v и для некоторых $m, n = 1, 2, 3$ существуют преобразовательные mn -потоки $f_v^{mn}(\tau)$ и $f_v^{nm}(\tau)$, внешние обменные потоки $d_{vu}^m(\tau)$ и $d_{uv}^m(\tau)$ для $u \in \{a_1, a_2\}$ и внутренние обменные mn -потоки $s_{uv}^{mn}(\tau)$ и $s_{vu}^{mn}(\tau)$, то в единице v для каждого $m = 1, 2, 3$ сумма интенсивностей всех входящих потоков за вычетом суммы всех накопительных долей от этих потоков равна сумме интенсивностей всех выходящих потоков.

Применяя это правило к единицам e , b , p и h , для $m=1,2,3$ получаем систему локальных балансовых уравнений нома (всего 12 уравнений - для каждого типа единицы и каждого m). Для примера выпишем 4 из них:

$$\begin{aligned}
(e_1) \quad & (1-\gamma_{a_1e}^1)d_{a_1e}^1 + (1-\gamma_{a_2e}^1)d_{a_2e}^1 + (1-\kappa_e^{11})f_e^{11} + (1-\kappa_e^{21})f_e^{21} + (1-\kappa_e^{31})f_e^{31} + \\
& \sum(\sum\{(1-\lambda_{fe}^{m1})s_{fe}^{m1} \mid f \in E \wedge \{e, f\} \in \overline{{}^2E}\} \mid m=1,2,3) = d_{ea_1}^1 + d_{ea_2}^1 + \sum(s_{eb}^{11} \mid b \in B_e) + f_e^{11} + \\
& + \sum(\sum\{s_{ef}^{1m} \mid f \in E \wedge \{e, f\} \in \overline{{}^2E}\} \mid m=1,2,3); \\
(b_2) \quad & (1-\gamma_{a_2b}^2)d_{a_2b}^2 + \sum\{(1-\lambda_{pb}^{22})s_{pb}^{22} \mid p \in P \wedge \exists e \in E(p = \varphi_e(b))\} + (1-\kappa_b^{12})f_b^{12} + \\
& + (1-\kappa_b^{22})f_b^{22} + (1-\kappa_b^{32})f_b^{32} + \sum(\sum\{(1-\lambda_{cb}^{m2})s_{cb}^{m2} \mid c \in B \wedge \{b, c\} \in \overline{{}^2B}\} \mid m=1,2,3) = \\
& = d_{ba_2}^2 + \sum\{s_{be}^{22} \mid e \in E \wedge b \in B_e\} + \sum(\sum\{s_{bc}^{2m} \mid c \in B \wedge \{b, c\} \in \overline{{}^2B}\} \mid m=1,2,3); \\
(p_2) \quad & (1-\gamma_{a_2p}^2)d_{a_2p}^2 + (1-\kappa_p^{12})f_p^{12} + (1-\kappa_p^{22})f_p^{22} + (1-\kappa_p^{32})f_p^{32} + \\
& + \sum(\sum\{(1-\lambda_{qp}^{22})s_{qp}^{22} \mid q \in P \wedge \{p, q\} \in \overline{{}^2P} \wedge p < q\} + \\
& + \sum(\sum\{(1-\lambda_{qp}^{m2})s_{qp}^{m2} \mid q \in P \wedge (p, q \in \text{Max}P \wedge \{p, q\} \in \overline{{}^2P})\} \mid m=1,2,3) = d_{pa_1}^2 + d_{pa_2}^2 + \\
& + \sum\{s_{pb}^{22} \mid b \in B \wedge \exists e \in E(p = \varphi_e(b))\} + \sum\{s_{ph}^{22} \mid h \in H \wedge \exists i \in I(h \in H_i \wedge (p, h) \in \zeta_i)\} + f_p^{22} \\
& + \sum\{s_{pq}^{22} \mid q \in P \wedge \{p, q\} \in \overline{{}^2P} \wedge q < p\} + \\
& + \sum(\sum\{s_{pq}^{2m} \mid q \in P \wedge (p, q \in \text{Max}P \wedge \{p, q\} \in \overline{{}^2P})\} \mid m=1,2,3); \\
(h_3) \quad & (1-\gamma_{a_2h}^3)d_{a_2h}^3 + (1-\kappa_h^{13})f_h^{13} + (1-\kappa_h^{23})f_h^{23} + (1-\kappa_h^{33})f_h^{33} + \\
& + \sum\{(1-\lambda_{kh}^{33})s_{kh}^{33} \mid k \in H \wedge \exists i \in I(\{k, h\} \in \overline{{}^2H_i} \wedge k < h)\} + \\
& + \sum(\sum\{(1-\lambda_{kh}^{m3})s_{kh}^{m3} \mid k \in H \wedge \exists i \in I(k, h \in \text{Max}H_i \wedge \{k, h\} \in \overline{{}^2H_i})\} \mid m=1,2,3) = \\
& = d_{ha_1}^3 + d_{ha_2}^3 + \sum\{s_{hp}^{33} \mid p \in P \wedge \forall i \in I(h \in H_i \Rightarrow (p, h) \in \zeta_i)\} + f_h^{33} + \\
& + \sum\{s_{hk}^{33} \mid k \in H \wedge \exists i \in I(\{k, h\} \in \overline{{}^2H_i} \wedge h < k)\} + \\
& + \sum(\sum\{s_{hk}^{3m} \mid k \in H \wedge \exists i \in I(k, h \in \text{Max}H_i \wedge \{k, h\} \in \overline{{}^2H_i})\} \mid m=1,2,3).
\end{aligned}$$

6 Правила ограничения достояний в некоторых единицах

Заданы ограничители (первичные параметры) $M_f^2(t)$ и $M_f^3(t)$, $M_h^1(t)$ и $M_h^2(t)$, $M_p^1(t)$ и $M_p^3(t)$ и правило ограничения:

$$\begin{aligned}
w_f^2(t) &\leq M_f^2(t) \text{ и } w_f^3(t) \leq M_f^3(t) \text{ для любого } t \in T \text{ и любого } f \in F; \\
w_h^1(t) &\leq M_h^1(t) \text{ и } w_h^2(t) \leq M_h^2(t) \text{ для любого } t \in T \text{ и любого } h \in H; \\
w_p^1(t) &\leq M_p^1(t) \text{ и } w_p^3(t) \leq M_p^3(t) \text{ для любого } t \in T \text{ и любого } p \in P.
\end{aligned}$$

7 Ном, его достояния и потоки

Математическую систему, состоящую из номовой системы $\hat{S} = (U, a_1, a_2; \sigma)$, интервала T и дополнительной структуры, задаваемой достояниями w_u^m , потоками f_u^{mn} , d_{uv}^m и s_{uv}^{mn} и коэффициентами γ_{uv}^m , κ_u^{mn} и λ_{uv}^{mn} , ограничителями M_u^m и удовлетворяющую правилам распределения потоков и ограничения достояний, назовем *номом* и будем обозначать через $N \equiv (U, a_1, a_2; \sigma; T; \nu)$.

Определим *совокупные m -достояния основных частей нома E , B , P и H на момент времени $t \in T$ для $m = 1, 2, 3$:*

- 1) $W_E^m(t) \equiv \sum (w_e^m(t) | e \in E)$;
- 2) $W_B^m(t) \equiv \sum (w_b^m(t) | b \in B)$;
- 3) $W_P^m(t) \equiv \sum (w_p^m(t) | p \in P)$;
- 4) $W_{H_i}^m(t) \equiv \sum (w_h^m(t) | h \in H_i)$;
- 5) $W_H^m(t) \equiv \sum (W_{H_i}^m(t) | i \in I)$.

Определим *совокупные потоки нома между основными частями нома и внешней средой:*

- 1) $d_{a_1E}^1(\tau) \equiv \sum (d_{a_1e}^1(\tau) | e \in E)$;
- 2) $d_{Ea_1}^1(\tau) \equiv \sum (d_{ea_1}^1(\tau) | e \in E)$;
- 3) $d_{a_2E}^m(\tau) \equiv \sum (d_{a_2e}^m(\tau) | e \in E) \quad (m = 1, 2, 3)$;
- 4) $d_{Ea_2}^m(\tau) \equiv \sum (d_{ea_2}^m(\tau) | e \in E) \quad (m = 1, 2, 3)$;
- 5) $d_{a_1B}^1(\tau) \equiv \sum (d_{a_1b}^1(\tau) | b \in B)$;
- 6) $d_{Ba_1}^1(\tau) \equiv \sum (d_{ba_1}^1(\tau) | b \in B)$;
- 7) $d_{a_2B}^m(\tau) \equiv \sum (d_{a_2b}^m(\tau) | b \in B) \quad (m = 1, 2, 3)$;
- 8) $d_{Ba_2}^m(\tau) \equiv \sum (d_{ba_2}^m(\tau) | b \in B) \quad (m = 1, 2, 3)$;
- 9) $d_{a_1P}^1(\tau) \equiv \sum (d_{a_1p}^1(\tau) | p \in P)$;
- 10) $d_{Pa_1}^1(\tau) \equiv \sum (d_{pa_1}^1(\tau) | p \in P)$;
- 11) $d_{a_2P}^m(\tau) \equiv \sum (d_{a_2p}^m(\tau) | p \in P) \quad (m = 1, 2, 3)$;
- 12) $d_{Pa_2}^m(\tau) \equiv \sum (d_{pa_2}^m(\tau) | p \in P) \quad (m = 1, 2, 3)$;
- 13) $d_{a_1H}^1(\tau) \equiv \sum (d_{a_1h}^1(\tau) | h \in H)$;
- 14) $d_{Ha_1}^3(\tau) \equiv \sum (d_{ha_1}^3(\tau) | h \in h)$;
- 15) $d_{a_2H}^m(\tau) \equiv \sum (d_{a_2h}^m(\tau) | h \in H) \quad (m = 1, 2, 3)$;
- 16) $d_{Ha_2}^m(\tau) \equiv \sum (d_{ha_2}^m(\tau) | h \in H) \quad (m = 1, 2, 3)$.

Определим *совокупные обменные потоки между основными частями нома:*

- 1) $s_{EB}^{11}(\tau) \equiv \sum (s_{eb}^{11}(\tau) | e \in E \wedge b \in B_e)$;
- 2) $s_{BE}^{mm}(\tau) \equiv \sum (s_{be}^{mm}(\tau) | e \in E \wedge b \in B_e) \quad \text{для } m = 2, 3$;
- 3) $s_{BP}^{11}(\tau) \equiv \sum (s_{bp}^{11}(\tau) | b \in B \wedge p \in P \wedge \exists e \in E (b \in B_e \wedge p = \varphi_e(b)))$;
- 4) $s_{PB}^{mm}(\tau) \equiv \sum (s_{pb}^{mm}(\tau) | b \in B \wedge p \in P \wedge \exists e \in E (b \in B_e \wedge p = \varphi_e(b))) \quad \text{для } m = 2, 3$;
- 5) $s_{PH}^{mm}(\tau) \equiv \sum (\sum (s_{ph}^{mm}(\tau) | p \in \tilde{P}_i \wedge h \in \tilde{H}_i \wedge (p, h) \in \zeta_i) | i \in I) \quad \text{для } m = 1, 2$;
- 6) $s_{HP}^{33}(\tau) \equiv \sum (\sum (s_{hp}^{33}(\tau) | p \in \tilde{P}_i \wedge h \in \tilde{H}_i \wedge (p, h) \in \zeta_i) | i \in I)$.

Определим *совокупные преобразовательные потоки в основных частях нома:*

- 1) $f_E^{m1}(\tau) \equiv \sum (f_e^{m1}(\tau) | e \in E) \quad (m = 1, 2, 3)$;
- 2) $f_B^{m2}(\tau) \equiv \sum (f_b^{m2}(\tau) | b \in B) \quad (m = 1, 2, 3)$;
- 3) $f_P^{m2}(\tau) \equiv \sum (f_p^{m2}(\tau) | p \in P) \quad (m = 1, 2, 3)$;
- 4) $f_H^{m3}(\tau) \equiv \sum (f_h^{m3}(\tau) | h \in H) \quad (m = 1, 2, 3)$.

Определим *совокупные накопительные t -доли основных частей нома от всех входящих потоков* (всего 12, для примера выпишем 4):

$$1) v_E^1 \equiv \sum (\gamma_{a_1 e}^1 d_{a_1 e}^1 + \gamma_{a_2 e}^1 d_{a_2 e}^1 + \kappa_e^{11} f_e^{11} + \kappa_e^{21} f_e^{21} + \kappa_e^{31} f_e^{31} | e \in E); \dots$$

$$5) v_B^2 \equiv \sum (\gamma_{a_2 b}^2 d_{a_2 b}^2 + \kappa_b^{12} f_b^{12} + \kappa_b^{22} f_b^{22} + \kappa_b^{32} f_b^{32} | b \in B) + \\ + \sum \{ \lambda_{pb}^{22} s_{pb}^{22} | b \in B \wedge p \in P \wedge \exists e \in E (p = \varphi_e(b)) \}; \dots$$

$$8) v_P^2 \equiv \sum (\gamma_{a_2 p}^2 d_{a_2 p}^2 + \kappa_p^{12} f_p^{12} + \kappa_p^{22} f_p^{22} + \kappa_p^{32} f_p^{32} | p \in P); \dots$$

$$12) v_H^3 \equiv \sum (\gamma_{a_2 h}^3 d_{a_2 h}^3 + \kappa_h^{13} f_h^{13} + \kappa_h^{23} f_h^{23} + \kappa_h^{33} f_h^{33} | h \in H).$$

8 Распределение совокупных потоков. Балансовые уравнения нома

Пусть N - фиксированный ном. Из правила распределения всех локальных потоков во всех единицах $v \in U$ выводится следующее *правило распределения всех совокупных t -потоков в основных частях E , B , P и H нома N* :

в каждой основной части для каждого $t = 1, 2, 3$ сумма интенсивностей всех входящих совокупных потоков за вычетом суммы всех накопительных долей от этих потоков равна сумме интенсивностей всех выходящих совокупных потоков.

Применяя это правило к основным частям E , B , P и H , для $t = 1, 2, 3$ мы получаем следующую *систему балансовых уравнений нома*:

$$(E_1) \quad d_{a_1 E}^1 + d_{a_2 E}^1 + f_E^{11} + f_E^{21} + f_E^{31} - v_E^1 = d_{Ea_1}^1 + d_{Ea_2}^1 + s_{EB}^{11} + f_E^{11}$$

$$(E_2) \quad d_{a_2 E}^2 + s_{BE}^{22} - v_E^2 = d_{Ea_2}^2 + f_E^{21}$$

$$(E_3) \quad d_{a_2 E}^3 + s_{BE}^{33} - v_E^3 = d_{Ea_2}^3 + f_E^{31}$$

$$(B_1) \quad d_{a_1 B}^1 + d_{a_2 B}^1 + s_{EB}^{11} - v_B^1 = d_{Ba_1}^1 + d_{Ba_2}^1 + s_{BE}^{11} + s_{BP}^{11} + f_B^{12}$$

$$(B_2) \quad d_{a_2 B}^2 + s_{PB}^{22} + f_B^{12} + f_B^{22} + f_B^{32} - v_B^2 = d_{Ba_2}^2 + s_{BE}^{22} + f_B^{22}$$

$$(B_3) \quad d_{a_2 B}^3 + s_{PB}^{33} - v_B^3 = d_{Ba_2}^3 + s_{BE}^{33} + f_B^{32}$$

$$(P_1) \quad d_{a_1 P}^1 + d_{a_2 P}^1 + s_{BP}^{11} - v_P^1 = d_{Pa_2}^1 + s_{PH}^{11} + f_P^{12}$$

$$(P_2) \quad d_{a_2 P}^2 + f_P^{12} + f_P^{22} + f_P^{32} - v_P^2 = d_{Pa_1}^2 + d_{Pa_2}^2 + s_{PF}^{22} + s_{PH}^{22} + f_P^{22}$$

$$(P_3) \quad d_{a_2 P}^3 + s_{HP}^{33} - v_P^3 = d_{Pa_2}^3 + s_{PF}^{33} + f_P^{32}$$

$$(H_1) \quad d_{a_1 H}^1 + d_{a_2 H}^1 + s_{PH}^{11} - v_H^1 = d_{Ha_2}^1 + f_H^{13}$$

$$(H_2) \quad d_{a_2 H}^2 + s_{PH}^{22} - v_H^2 = d_{Ha_2}^2 + f_H^{23}$$

$$(H_3) \quad d_{a_2 H}^3 + f_H^{13} + f_H^{23} + f_H^{33} - v_H^3 = d_{Ha_1}^3 + d_{Ha_2}^3 + s_{HP}^{33} + f_H^{33}.$$

9 Условия подкрепления. Импульсные уравнения естественного нома

Ном будем называть удовлетворяющим *локальным условиям подкрепления*, если заданы первичные параметры $\varepsilon_{be}(\tau)$, $\varepsilon_{pb}(\tau)$, $\varepsilon_{qp}(\tau)$, $\varepsilon_{ph}(\tau)$ и $\varepsilon_{hk}(\tau)$ такие, что:

$$1) \quad s_{be}^{33}(\tau) = \varepsilon_{be}(\tau) s_{be}^{22}(\tau) \text{ для любых } e \in E \text{ и } b \in B_e;$$

- 2) $s_{pb}^{33}(\tau) = \varepsilon_{pb}(\tau)s_{pb}^{22}(\tau)$ для любых $b \in B$ и $p \in P$, для которых существует $e \in E$ такое, что $p = \varphi_e(b)$;
- 3) $s_{qp}^{33}(\tau) = \varepsilon_{qp}(\tau)s_{qp}^{22}(\tau)$ для любых $p, q \in P$ таких, что $\{p, q\} \in \overline{{}^2P}$ и $p < q$;
- 4) $s_{ph}^{11}(\tau) = \varepsilon_{ph}(\tau)s_{ph}^{22}(\tau)$ для любого $i \in I$ и любых $p \in P$ и $h \in H_i$ таких, что $(p, h) \in \zeta_i$;
- 5) $s_{kh}^{11}(\tau) = \varepsilon_{kh}(\tau)s_{kh}^{22}(\tau)$ для любого $i \in I$ и любых $k, h \in H_i$ таких, что $\{h, k\} \in \overline{{}^2H_i}$ и $h < k$.

В таком номе все распорядительные потоки из распорядительных единиц подкрепляются соответствующим содержательным или обеспечительным потоками из этих единиц.

Пусть для $u, v \in U$ имеется mm -поток s_{uv}^{mm} . Произведение $w_u^m s_{uv}^{mm}$ будем называть *m-импульсом единицы u на единицу v* . Импульсы b на e , p на b , q на p (при $p < q$), p на h и h на k (при $k < h$) будем называть *воздействующими*, остальные импульсы будем называть *ответными*.

Ном с локальными условиями подкрепления будем называть *локально естественным*, если для него справедливо следующее *правило передачи импульсов в единицах*: *сумма всех воздействующих импульсов единицы u на единицу v равна сумме всех ответных импульсов единицы v на единицу u* .

Формально это правило записывается в виде следующих *локальных импульсных уравнений*:

$$\begin{aligned}
 (b \downarrow e) \quad & w_b^2 s_{be}^{22} + w_b^3 s_{be}^{33} = w_e^1 s_{eb}^{11} \text{ для } b \in B_e; \\
 (p \downarrow b) \quad & w_p^2 s_{pb}^{22} + w_p^3 s_{pb}^{33} = w_b^1 s_{bp}^{11} \text{ для } p = \varphi_e(b); \\
 (q \downarrow p) \quad & w_q^2 s_{qp}^{22} + w_q^3 s_{qp}^{33} = w_p^1 s_{pq}^{11} \text{ для } p < q \text{ в } P; \\
 (p \downarrow h) \quad & w_p^1 s_{ph}^{11} + w_p^2 s_{ph}^{22} = w_h^3 s_{hp}^{33} \text{ для } (p, h) \in \zeta_i; \\
 (k \downarrow h) \quad & w_k^1 s_{kh}^{11} + w_k^2 s_{kh}^{22} = w_h^3 s_{hk}^{33} \text{ для } h < k \text{ в } H_i.
 \end{aligned}$$

Заключение

В данной работе рассмотрена только модель нома с *постоянным* строением, или, иначе, нома в *квазиустойчивом состоянии*. Большой интерес представляют такие вопросы, как *ном с переменным строением*, включение нома в качестве *поднома* в другой ном, *взаимодействие* номов. Эти вопросы предполагается рассмотреть в последующих работах.

Список литературы

- [1] Атаманчук Г.В. Теория государственного управления. - М.: ОМЕГА-Л, 2004. – 584 с.
- [2] Мильнер Б.З. Теория организации. - М.: ИНФА-М, 2003. – 558 с.
- [3] Петров А.А., Поспелов И.Г., Шананин А.А. Опыт математического моделирования экономики. – М.: Энергоатомиздат, 1996. – 544 с.
- [4] Самарский А.А., Михайлов А.П. Математическое моделирование. - М.: Физматлит, 2002. – 320 с.

ЗАДАЧА СИНТЕЗА СТРУКТУРЫ ЛОГИСТИЧЕСКОГО ЦЕНТРА

С.В. Архангельский, Н.А. Симон

Научно-производственный центр информационных и транспортных систем (НПЦ ИНФОТРАНС)
443001, Самара, ул. Полевая, 47
archangelski@infotrans-logistic.ru, nick@infotrans-logistic.ru
тел. +7 (8462) 32-53-17, 33-37-79, факс: +7 (8462) 32-31-66

Ключевые слова: *общетранспортные узлы, логистический центр, складская инфраструктура, оптимизация, модель, система массового обслуживания*

Abstract

The article has been set a problem of structure synthesis for logistical center. Structure and cargo handling in center's warehouse have been described. Analytical model has been developed for this warehouse.

Введение

В настоящее время в регионе Самарской области отсутствует единая информационная инфраструктура транспортно-экспедиционного обслуживания. Все транспортные системы технологически и организационно разобщены. Разобщенность видов транспорта не позволяет создать единое поле транспортно-логистических услуг. Каждый вид транспорта может перевозить грузы от пункта нахождения до назначенного пункта выгрузки. Однако подавляющее число грузов, прежде чем он поступит к потребителю, перегружается в среднем 3-4 раза. Как правило, это происходит в общетранспортных узлах (ОТУ). Именно в ОТУ на "стыках" различных видов транспорта общего пользования находятся основные резервы ускорения продвижения грузов, сокращения транспортных издержек, повышение качества транспортных услуг и благоприятная среда построения транспортных логистических цепей. В результате проведенного ранее анализа возможных схем организации грузопереработки была предложена схема с использованием логистического центра (ЛЦ), располагающегося в ОТУ [1].

Но до настоящего времени рассматриваемая предметная область характеризуется наличием неустойчивого определения понятия "логистический центр". Мы будем использовать следующее определение: "Логистический центр - это самостоятельный хозяйственный объект, оказывающий логистические услуги (перевозка, перегрузки, складирование, сортировка и комплектация грузов и т.д.) и реализующий тем самым функции снабжения и дистрибуции в определенном регионе" [2]. ЛЦ имеет, как правило, следующую организационно-функциональную структуру:

- 1) Подсистема организационного обеспечения (участок приема заказов; маркетинговая деятельность; расчет с поставщиками, потребителями и транспортными предприятиями). Она предназначена для продвижения на рынке услуг ЛЦ, создания и совершенствования договорной основы деятельности участников логистического процесса, отработки заказов и расчетов с клиентами и партнерами;
- 2) Подсистема информационного обеспечения (информационная система ЛЦ). Является главной организующим звеном ЛЦ, обеспечивает эффективное функционирование всех его подразделений, а также обмен информацией с клиентами и партнерами;
- 3) Автоматизированный склад. Он является основным производственным звеном логистического центра. Здесь осуществляются все технологические операции с грузами, проходящими через ЛЦ (погрузочно-разгрузочные работы, контроль, отбраковка, подготовка к

хранению, складирование, коммиссионирование, комплектование партий и отправка грузов).

- 4) Подсистема сервисного обеспечения (участок подготовки маршрутов движения и планирования загрузки транспорта, участок информирования о состоянии груза).

1 Постановка задачи

Задача оснащение склада ЛЦ соответствующим оборудованием и привлечение специалистов на склад являются одними из самых затратных статей бюджета ЛЦ. Следовательно, задача синтеза структуры ЛЦ фактически сводится к синтезу структуры автоматизированного склада ЛЦ. Для повышения эффективности функционирования ЛЦ необходимо выбрать такое количество оборудования и персонала или организовать технологический процесс на складе ЛЦ таким образом, чтобы загрузка ресурсов склада ЛЦ была максимальной, но при этом склад должен справляться с поставленными перед ним задачами. Поэтому, при проектировании склада ЛЦ необходимо использовать следующие критерии оценки эффективности работы склада ЛЦ: объем предоставляемых клиентам услуг, затраты на обработку заказа, длительность обработки заказов, коэффициент использования ресурсов ЛЦ.

Целевая функция оптимизации структуры и процесса функционирования ЛЦ является многопараметрической и многокритериальной. Она имеет следующий общий вид:

$$(1) \quad E(I, O, S, G, U) = [F_1(I, O, S, G, U) \rightarrow \text{extr}, \dots, F_m(I, O, S, G, U) \rightarrow \text{extr}],$$

где E - вектор критериев оценки эффективности работы ЛЦ; $F_1 \dots F_m$ - критерии оценки эффективности работы ЛЦ; I - вектор параметров, описывающих структуру входящего потока транспорта; O - вектор параметров, описывающих структуру исходящего потока транспорта; S - вектор параметров, характеризующих структуру и оснащенность ЛЦ; G - вектор параметров, описывающих график работы ЛЦ и организацию обработки заказов; U - параметры, характеризующие объем услуг, предоставляемых ЛЦ своим клиентам.

Объем предоставляемых услуг формируется на основе пожеланий заказчиков, территориального размещения ЛЦ, его финансовых возможностей и стратегических планов. Поэтому будем считать, что объем предоставляемых услуг является неизменной и оптимальной для ЛЦ и его клиентов величиной. В аналитической форме остальные критерии формируют следующую систему целевых функций:

$$(2) \quad \left\{ \begin{array}{l} C_{in} n_{in}(S, G) + \sum_{j=1}^h (C_{y,j} + C_{y,d} W_j) + \\ \sum_{k=1}^c (C_{ic} L_k^{dc} T_{dc,k}(S, G) + C_{ic} L_k^{da} T_{da,k}(S, G)) \rightarrow \min \\ \sum_{k=1}^c \left| l_k T_k^{ec} - \sum_{i=1}^p L_k^i T_{ik}(S, G) \right| \rightarrow \min \\ \sum_{j=1}^h |K_j^{\max} - K_j(S, G)| \rightarrow \min \end{array} \right. ,$$

при ограничениях:

$$(3) \quad \left\{ \begin{array}{l} K_j^{\max} \leq 1, j \in 1..h \\ K_j^{\max} \geq 0, j \in 1..h \\ T_{j,k}(S, G) \leq T_j^{\max}, j \in 1..p, k \in 1..c \end{array} \right. ,$$

где: p – количество этапов обработки заказов; c – количество заказов потребителей, обслуженных за рассматриваемый период; h – количество ресурсов (площадей и объемов складов, единиц складского оборудования, участков разгрузки и погрузки, персонала и т.д.); C_{in} – штраф за отказ обработки заказа; $n_{in}(S, G)$ – количество транспортных средств, получивших отказ в рассматриваемый период; $C_{y,j}$ – стоимость эксплуатации j -го ресурса в течении рассматриваемого периода; $C_{y,d\ddot{e}}$ – стоимость эксплуатации единицы площади склада; W_j – площадь, занимаемая j -м ресурсом; C_{ic} – штраф за ожидание транспортного средства; $L_k^{d\dot{c}}$ – признак разгрузки k -го заказа; $T_{d\dot{c},k}(S, G)$ – длительность разгрузки k -го заказа; $L_k^{d\ddot{a}}$ – признак погрузки k -го заказа; $T_{d\ddot{a},k}(S, G)$ – длительность погрузки k -го заказа; l_k – важность показателя длительности обработки k -го заказа для ЛЦ; $T_k^{\ddot{e}c}$ – требуемая длительность обработки заказа; L_k^i – признак прохождения i -го этапа k -м заказом; $T_{ik}(S, G)$ – длительность i -го этапа обработки k -го заказа; $K_j(S, G)$ – загрузка j -го ресурса ЛЦ; K_j^{\max} – максимально необходимая загрузка j -го ресурса; T_i^{\max} – максимально возможная длительность обработки заказа на j -м этапе.

Предположим, что между параметрами ЛЦ и критериями эффективности существует детерминированная зависимость. При синтезе ЛЦ происходит выбор оборудования и персонала для каждого участка между доступным на рынке оборудованием и персоналом. Каждый тип оборудования и квалификация персонала имеет свою производительность, стоимость создания (покупки/привлечения) и эксплуатации. Следовательно, на этапе определения типа и количества используемых ресурсов параметры задачи синтеза ЛЦ являются дискретными. Кроме того, функции $n_{in}(S, G)$, $T_{d\dot{c},k}(S, G)$, $T_{d\ddot{a},k}(S, G)$, $T_{ik}(S, G)$ и $K_j(S, G)$ в общем случае являются нелинейными. Из этого можно сделать вывод о том, что задача синтеза структуры ЛЦ является оптимизационной задачей нелинейного программирования (в случае свертки векторного критерия в скалярный) или многокритериального программирования.

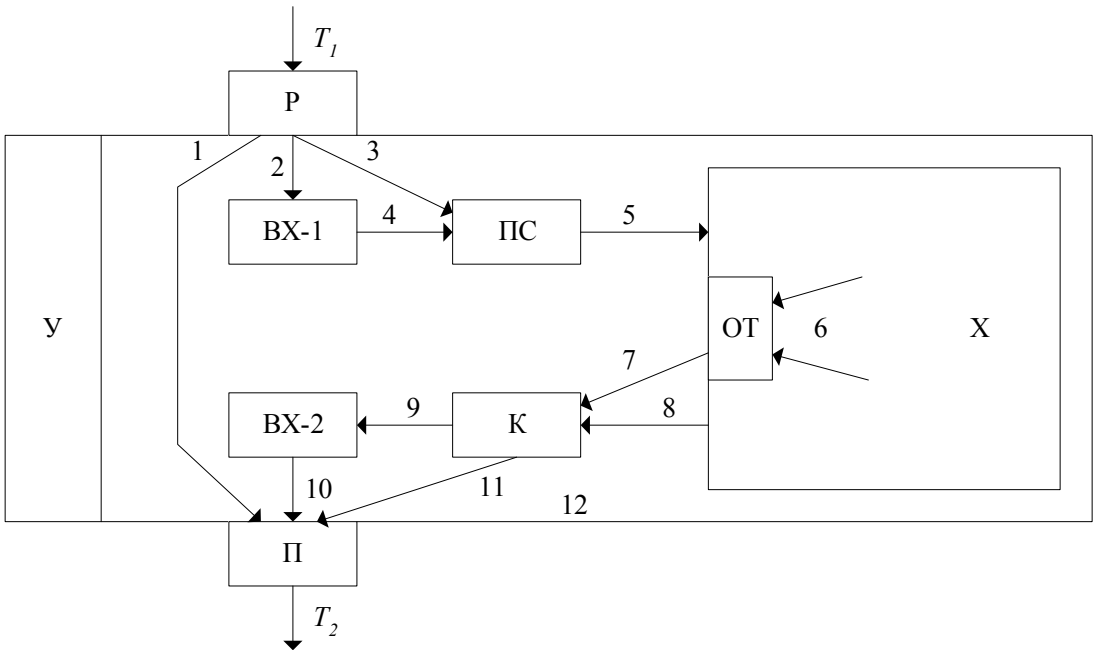
2 Декомпозиция структуры и технологического процесса склада ЛЦ

Так как эксперименты с уже созданным ЛЦ практически невозможны (в силу их большой стоимости и длительности), то для решения этой задачи необходимо создать и проанализировать некоторую модель склада ЛЦ. Модель склада ЛЦ строится на основе описания структуры и технологического процесса склада ЛЦ общего вида, созданного в результате анализа структуры и процесса работы нескольких складов. Оно включает в себя описание независимых участков склада; объектов, обрабатываемых на складе; устройств обслуживания, используемых на складе; алгоритмов работы склада и т.д. При описании структуры и работы складского хозяйства ЛЦ был использован метод декомпозиции: склад был разделен на составные элементы, в технологическом процессе были выделены функционально законченные операции.

Во-первых, склад был представлен как техническая система, состоящая из элементов (называемых технологическими участками), предложенных в [3]. На рисунке 1 показана структурная схема склада, с направлениями основных грузопотоков и технологическими взаимосвязями участков. Во-вторых, весь технологический процесс на складе ЛЦ был разбит на следующие функционально законченные операции: разгрузка и первичная приемка; приемка его по количеству и качеству; складирование и хранение; отборка и комиссионирование; комплектация партий отправки; отгрузка; внутрискладская транспортировка.

После определения основных объектов склада ЛЦ было создано описание технологического процесса склада. Описание и анализ этого процесса проводилось "сверху - вниз". Перво-

начально был описан процесс в целом, а затем все функционально законченные операции технологического процесса.



T_1 — транспорт прибытия; T_2 — транспорт отправления; P — участок разгрузки; VX-1 — участок временного хранения прибывающих грузов; PC — участок приема и сортировки поступающих грузов; X — участок хранения; OT — участок отборки грузов; K — участок комплектации грузов; VX-2 - участок временного хранения выдаваемых грузов; П — участок погрузки грузов; У — подсистема управления складом; 1—12 —складские грузопотоки.

Рисунок 1 – Структурная схема склада ЛЦ.

В результате описания процесса грузопереработки на складе ЛЦ общего вида в целом была построена таблица 1 указывающая, какие участки склада необходимо моделировать для анализа каждой операции технологического процесса склада.

Таблица 1 – Взаимосвязь между технологическими участками и операциями склада ЛЦ.

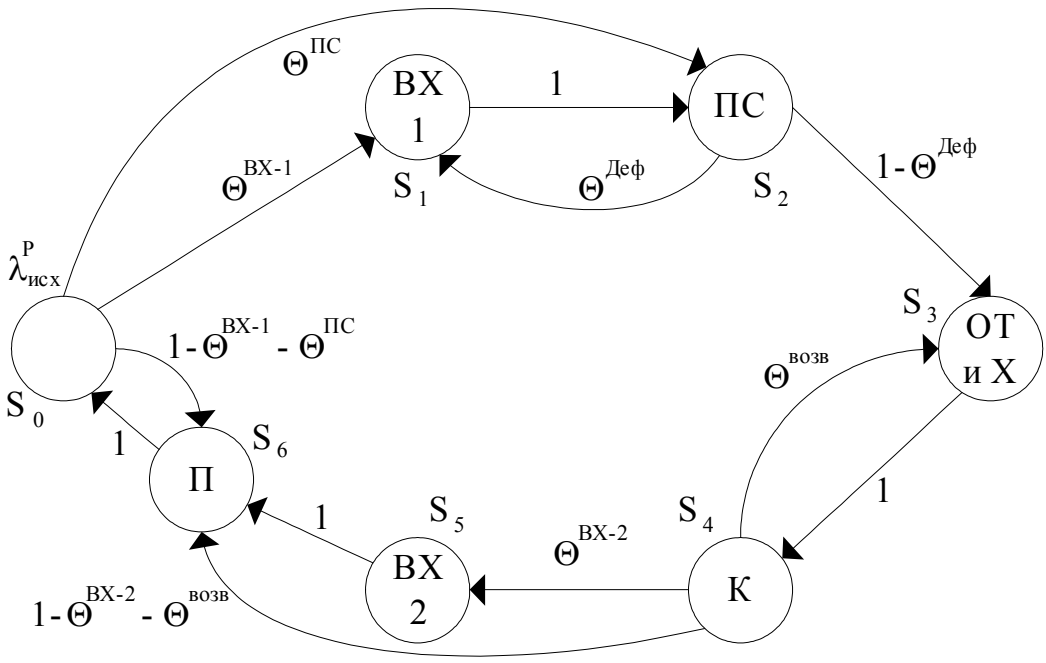
Операции	Участки склада							
	P	VX-1	PC	X	OT	K	VX-2	П
Разгрузка и первичная приемка	+							
Приемка по количеству и качеству			+					
Складирование и хранение			+	+				
Отборка и комиссионирование				+	+	+		
Комплектация партий отправки						+		
Отгрузка								+
Внутрискладская транспортировка	+	+	+	+	+	+	+	+

3 Аналитическая модель склада ЛЦ

Полученное описание структуры и технологического процесса склада позволяет сделать следующий вывод: определение количественных характеристик склада можно провести с помощью семейства аналитических моделей в терминах теории массового обслуживания. Не-

смотря на то, что использование этих моделей связано с некоторыми допущениями (например, входящий поток транспортных средств будем считать простейшим), они позволят быстро определить необходимое количество и номенклатуру оборудования и персонала склада ЛЦ для заданных объем грузопотоков.

Но склад ЛЦ, рассматриваемый на уровне участков, невозможно представить в качестве сети СМО, так как различные участки склада ЛЦ работают с разными видами заявок на обслуживание (транспортные средства (ТС), транспортные (ТГЕ) и складские (СГЕ) грузовые единицы и т.д.). Поэтому, для аналитической модели было сделано предположение о том, что после разгрузки ТГЕ не разделяются и проходят весь технологический цикл в виде единой заявки. В этом случае, складское хозяйство можно представить, как совокупность двух многоканальных СМО (участки Р и П, работающие с ТС) и разомкнутой сети СМО, описывающей остальные участки. Граф передач для данной сети, который также содержит узел П для оценки входящего потока ТГЕ на участок П, представлен на рисунке 2.



$\lambda_P^{исх}$ - интенсивность потока с участка Р, Θ_{BX-1} - вероятность перемещения ТГЕ на участок ВХ-1, $\Theta_{ПС}$ - вероятность перемещения ТГЕ на участок ПС, $\Theta_{Деф}$ - вероятность обнаружения дефектной ТГЕ, $\Theta_{возв}$ - вероятность возвращения СГЕ с К, Θ_{BX-2} - вероятность перемещения ТГЕ на участок ВХ-2.

Рисунок 2 – Граф передач сети СМО склада ЛЦ.

На основании данного графа были определены интенсивности потока заявок на каждый участок склада, которые были использованы для моделирования участков по отдельности. Для них были созданы разомкнутые сети СМО, описывающие работу каждого участка склада ЛЦ.

В данной модели склада ЛЦ используются усредненные параметры входящего и исходящего потока. В результате моделирования получаем усредненные результаты для стационарного режима работы склада ЛЦ. Кроме того, модель предназначена только для определения необходимого оборудования ЛЦ.

Параметры входящего и исходящего потоков транспорта, количество видов транспорта и вероятности передач заявок между участками склада являются неизменными. Варьируемым параметром данной модели является только вектор S , который включает в себя количество и тип используемого оборудования. Тип E выбирается из конечного количества доступных на рынке решений. Он однозначно определяет следующие параметры: производительность μ (для мест хранения (МХ) производительность не зависит от E , а равна среднему времени хранения на соответствующем участке), стоимость эксплуатации C_y и создания $C_{на}$, занимаемая площадь W (для МХ занимаемая площадь рассчитывается исходя из максимальной высоты подъема обслуживающих ПТУ и общего количества МХ на участке).

Полученная модель характеризуется большим количеством варьируемых параметров (при работе с двумя видами транспорта модель имеет 34 параметра) и нелинейные критерии оценки эффективности, что затрудняет поиск оптимальных вариантов структуры склада ЛЦ. Для поиска оптимальной структуры за конечное машинное время можно использовать следующие пути:

- Поиск оптимальной структуры каждого участка склада по отдельности. Но в этом случае необходимо дополнительная информация о распределении временных нормативов между участками;
- Использовать участок в качестве минимального уровня моделирования. Но в этом случае снизится точность моделирования, и выбор будет осуществляться только среди производителей, предлагающих комплексную автоматизацию участков склада.

Наиболее плодотворным является первый подход. При этом исходная задача разбивается на ряд подзадач (по числу участков склада ЛЦ) с меньшим числом параметров. Временные нормативы для данных моделей можно определить на основании статистических данных о работе крупных складов.

Заключение

В результате анализа работы нескольких складов было создано описание структуры и технологического процесса склада ЛЦ общего вида. Оно позволило сделать вывод о целесообразности использования для моделирования склада ЛЦ аналитических моделей в терминах теории массового обслуживания.

На основании описания складского хозяйства ЛЦ была разработана аналитическая модель складского хозяйства ЛЦ, предназначенная для его синтеза. Но в силу того, что аппарат СМО позволяет определить только усредненные характеристики СМО для потока заявок с усредненными параметрами, для повышения точности моделирования планируется использовать также имитационную модель. В этом случае, результат, полученный при проектировании структуры ЛЦ с помощью аналитической модели, будет являться отправной точкой для имитационного моделирования. Кроме этого, имитационная модель позволит уточнить полученные результаты, определить пространственное расположение оборудования и использовать в модели опытные данные о поступлении/отправлении транспорта, работе персонала и оборудования.

Список литературы

- [1] Архангельский С.В., Ефремов В.А., Симон Н.А. Методологические аспекты проектирования региональных логистических систем // Проблемы управления и моделирования в сложных системах: Труды IV международной конференции (Самара, 17-23 июня 2002 г.). – Самара: СНЦ РАН, 2002.
- [2] Фиялковский Я. Опыт проектирования логистических центров в Польше // Логинфо № 10, 2001. – С. 34-37.
- [3] Маликов О.Б. Деловая логистика. - СПб.: Политехника, 2003.

МОДЕЛЬ РЕГУЛЯЦИИ БИОСИНТЕЗА

К.Ю. Горбунов, В.А. Любецкий

Институт проблем передачи информации РАН
127994, Москва, Б. Каретный пер, 19
gorbunov@iitp.ru, lyubetsk@iitp.ru
тел: +7 (095) 413-46-43, факс +7 (095) 209-05-79

Ключевые слова: биоинформатика, регуляция биосинтеза, аттенуаторная регуляция, репрессорная регуляция

Abstract

We propose a methodological approach based on the generalized automata theory (directed graphs with potentially infinite number of marks). This theory can be used to model metabolite biosynthesis in bacterial culture. Results of testing of our algorithm are given.

Введение

Широко известна задача моделирования биосинтеза веществ в бактериальной культуре и, в частности, построения модели регуляции экспрессии генов, см., например, [1, 2]. Попытки такого моделирования, основанные на аппаратах дифференциальных уравнений или случайных процессов, сталкиваются с известными вычислительными трудностями, а также с трудностями интерпретации результатов. Последние связаны с косвенным характером описания биологических процессов при использовании такого математического аппарата. Желательно развитие более прямого и вычислительно более прозрачного аппарата для решения этой задачи.

Мы предлагаем подход к указанной задаче, основанный на обобщенной теории автоматов (ориентированный граф с потенциально бесконечным числом меток). На этом языке описываются процессы транскрипции и трансляции – в этом сообщении в части репрессорной и аттенуаторной регуляций экспрессии оперона и биосинтеза аминокислот. А именно, предлагается выделить конечное число вершин графа – «состояний системы»: из одного состояния возможен переход в одно из нескольких других в зависимости от, например, концентрации веществ и значений некоторых случайных величин. Скажем, в зависимости от концентрации триптофана из «паузно-кодонного» состояния (когда РНК-полимераза находится в последнем такте паузы, а рибосома – на последнем регуляторном кодоне) возможен переход в «послепаузно-кодонное» состояние (когда рибосома осталась на том же регуляторном кодоне) или переход в «послепаузно-послекодонное» состояние (когда рибосома сошла с регуляторных кодонов). При этом вершинам графа на каждом такте динамически приписываются числа, характеризующие: долю бактерий из культуры, находящихся в данном состоянии, и некоторых средние показатели культуры. На основе такого подхода нами, в частности, компьютерно моделировалась регуляция биосинтеза триптофана. Были получены графики, которые показывают, например, что репрессорная регуляция быстрее реагирует на концентрацию триптофана, чем аттенуаторная (факт экспериментально известный).

1 Описание модели

Множество возможных состояний представим множеством вершин ориентированного графа, в котором каждое направленное ребро ведёт из состояния в другое возможное состояние. Каждой вершине приписано рациональное число (от 0 до 100), которое интерпретируется как процент бактерий в культуре, у которых фиксированный оперон (ортологичный ряд оперо-

нов) находится в состоянии, соответствующем этой вершине. Тактовая динамика на этом графе состоит в изменении этого (процентного) распределения и некоторых средних характеристик культуры.

Рассмотрим наш подход на примере синтеза триптофана. Рассматривается динамика изменения концентраций ферментов синтеза триптофана, самого триптофана и активного репрессора. Для простоты мы говорим сейчас об одном ферменте и одном репрессоре, хотя модель легко позволяет рассматривать несколько ферментов и репрессоров/активаторов. Конечно, в разных клетках концентрации веществ в каждый момент различаются. Мы оперируем с их средними значениями (так же, как в [1, 2] и многих работах), принимая, что зависимости между средними значениями (по крайней мере, в среднем) соответствуют зависимостям между концентрациями этих веществ в отдельной клетке. Если мы хотим учитывать неоднородность концентраций, то можно отражать в состоянии оперона степень концентрации каждого из веществ. Вероятность посадки РНК-полимеразы (за один такт времени) зависит от текущей концентрации активного репрессора; другие вероятности зависят также от различных концентраций. Регуляция производства триптофана выражается в следующем. С ростом концентрации триптофана: растёт вероятность сдвига рибосомы с регуляторного кодона; усиливается ингибирование фермента триптофаном; падает количество триптофана, производимого 1% фермента; растёт концентрация активного репрессора. Вид этих зависимостей подбирался на основе известных биологических наблюдений и из общих соображений о характере биосинтеза. Учитывались также процессы утилизации триптофана, его транспорта из внешней среды, распада фермента и роста бактериальной культуры.

Наша модель процесса биосинтеза триптофана включает: множество состояний бактерий и функцию (возможно, вероятностную), которая по процентному распределению бактерий по состояниям и текущей концентрации фермента, триптофана и активного репрессора выдаёт новое процентное распределение и новую концентрацию этих веществ (для следующего такта).

Вводится конечное состояние, интерпретируемое как «выход из игры», на которое процентное распределение не распространяется. Количество попавших в конечное состояние бактерий сразу возмещается прибавкой бактерий в начальном состоянии (оперон свободен). Таким образом, новое процентное распределение бактерий однозначно определяется по старому с помощью набора функций f_q , где для каждого состояния q функция $f_q(q')$ указывает, как (процентно) распределяются по состояниям q' бактерии, находившиеся в состоянии q , по окончании такта.

Для простоты считаем, что успех или неуспех синтеза определяется тем, кто (рибосома или РНК-полимераза) раньше пересечёт свою «линию финиша», расположенную у рибосомы на заданном расстоянии от группы регуляторных кодонов (считаем их расположенными рядом; далее кодоном будем называть регуляторный кодон), а у РНК-полимеразы – на заданном расстоянии от места паузы. Исключением считаем случай, когда в момент пересечения полимеразой «линии финиша» рибосома вообще еще не села – этот случай относим к неуспеху. Нами предложены различные варианты функций f_q , а также функции, определяющие концентрации фермента, триптофана и репрессора на следующем такте по их концентрациям на предыдущем такте.

2 Результаты моделирования и выводы

Принималось, что скорости РНК-полимеразы и рибосомы постоянны и равны между собой, время паузы постоянно. За единицу времени бралось время транскрипции полимеразой одного нуклеотида. При этих условиях многие параметры известны из биологических данных или приближённо рассчитываются. Численное варьирование параметров, значение которых не вполне ясно, показало, что графики итоговых зависимостей (по крайней мере, на качественном уровне) мало от них зависят.

В начальном моменте все бактерии находились в состоянии q_0 , концентрации фермента, триптофана и репрессора брались нулевыми. В первые 135 тактов (когда производство фермента ещё не налажилось) концентрация фермента была нулевой, концентрация триптофана – очень близкой к нулю, концентрация репрессора колебалась около 0.1. Затем концентрация фермента стала быстро расти и к такту 139 достигла 0.66, а чуть позже (к такту 153) и концентрация триптофана достигла 1.68, что считается почти критическим избытком. В это время была велика и концентрация репрессора, равная 1.75. Затем все концентрации стали падать и к такту 271 концентрации фермента и триптофана достигли примерно значения 0.07, что соответствует недостатку триптофана.

Такие большие колебания были характерны лишь для начального отрезка времени, когда система регулировки синтеза ещё не стабилизировалась. В дальнейшем колебания затухают. Затухание происходит не очень быстро: в середине второй тысячи тактов всё ещё заметны колебания концентрации фермента примерно от 0.08 до 0.13 и почти синхронно с ней колеблется концентрация триптофана (от 0.08 до 0.14) и концентрация репрессора (от 0.86 до 1.18). Эти колебания соответствуют естественному биологическому циклу. Отметим, что, как показали испытания, предельные значения концентраций не зависят от начального распределения бактерий по состояниям и от начальных концентраций.

Для сопоставления аттенуаторной и репрессорной систем регуляции проводилось ещё два испытания. В первом из них была резко усилена роль репрессора (соответствующие параметры были заданы равными 0.1, т.е. в два раза меньше стандартной величины). В результате затухание колебаний резко ускорилось: в середине второй тысячи тактов цикл практически исчез: концентрация фермента стабилизировалась на уровне 0.11, а концентрация триптофана – на уровне 0.12. Во втором испытании, наоборот, роль репрессорной системы регуляции была ослаблена (соответствующие параметры равнялись 0.5). В результате колебания стали сильнее, а их затухание замедлилось: в середине второй тысячи тактов концентрация фермента колебалась от 0.06 до 0.14, а концентрация триптофана – от 0.06 до 0.17. Таким образом, можно сделать вывод, что репрессорная система регуляции действует более оперативно и быстрее реагирует на недостаток или избыток триптофана, чем аттенуаторная. Этот вывод совпадает с известным положением о большей чувствительности репрессорной системы по сравнению с аттенуаторной.

Чтобы приблизить модель к биологической реальности в неё также вводился элемент случайности. А именно, поскольку в реальности потребность клетки в триптофане непостоянна, отражающий эту потребность параметр колебался случайным образом около своего среднего значения. При этом все описанные закономерности сохранились, но колебания стали более хаотическими и перестали затухать в пределе.

Список литературы

- [1] Santillam M., Mackey M.C. Dynamic regulation of the tryptofan operon: A modeling study and comparison with experimental data // Proc. Natl. Acad. Sci. USA, v. 98, is. 4, 2001, p.p. 1364-1369.
- [2] Elf J., Berg O.G., Ehrenberg M. Comparison of Repressor and Transcriptional Attenuator Systems for Control of Amino Acid Biosynthetic Operons // J. Mol. Biol., v. 313, 2001, p.p. 941-954.

ГЕОМЕТРИЧЕСКИЙ МЕТОД ПОИСКА КЛИКИ В ГРАФЕ И ЕГО ПРИМЕНЕНИЕ ДЛЯ ВЫДЕЛЕНИЯ СИГНАЛА

В.А. Любецкий, А.В. Селиверстов

Институт проблем передачи информации, Российская академия наук
127994, Москва, Большой Каретный переулок, 19
lyubetsk@iitp.ru, slvstv@iitp.ru
тел: +7 (095) 413-46-43, факс +7 (095) 209-05-79

Ключевые слова: *многодольный граф, клика, политоп, выделение сигнала.*

Abstract

We consider finding and counting n -clique in any n -partite graph with two vertices in each part. This task can be reduced to linear programming.

Введение

Нами предложен существенно новый алгоритм поиска клики в многодольном графе с двумя вершинами в каждой доле, который сводит эту сложную задачу к задаче линейного программирования (вторая из них представляется много более простой). Известно, что поиск n -клики в n -дольном графе с двумя вершинами в каждой доле сводится за полиномиальное время к поиску модели пропозициональной конъюнктивной нормальной формы с двумя литералами в каждой дизъюнкции (2КНФ), которая в свою очередь может быть найдена за полиномиальное время, [1]. Предлагаемый нами подход позволяет:

- Решать исходную задачу поиска клики с помощью симплекс-метода,
- Получать при этом дополнительную информацию, например оценку сверху числа n -клик.

1 Постановка задачи

Алгоритмы поиска клики позволяют решать задачи выравнивания последовательностей, т.е. поиска набора похожих слов, по одному слову в каждой из n пар. Такой набор слов часто называют сигналом. Поиск сигнала интересен для решения задач биоинформатики, например, для выравнивания нуклеотидных последовательностей. В этой задаче по n данным последовательностям в алфавите $\{a, c, g, t\}$ строится n -дольный граф G , вершинами которого служат слова из этих последовательностей фиксированной длины. Две вершины соединены ребром в графе G , если они являются словами из разных последовательностей и похожи друг на друга больше некоторого фиксированного порога. Системе попарно похожих слов (по одному слову в каждой из n последовательностей) соответствует n -клика в графе G . Если из некоторых сообщений в каждой последовательности выбрано по два слова, то каждая доля графа G содержит по две вершины.

К поиску клики в многодольном графе сводится и другая задача. Пусть известна пространственная структура некоторого белка большого размера. Требуется найти место посадки небольшой молекулы-лиганда. Эта задача имеет простую механическую модель: найти участок поверхности, к которому можно достаточно плотно прижать данный шаблон.

Напомним основные понятия. Граф G называется n -дольным, если множество вершин графа разбито на n долей так, что концы любого ребра принадлежат разным долям. Многодольный граф G называется *полным многодольным*, если каждые две вершины из разных долей соединены ребром. Полный многодольный граф является полным графом тогда и только тогда,

когда каждая доля состоит из одной вершины. Полный подграф, имеющий n вершин, называется n -кликкой. Детали можно найти в [2].

2 Политопы и графы

Ниже индексы p, q, r равны 1 или 2, а индексы i, j, k пробегает значения от 1 до n . Для целого числа n определим политоп P_n в $4n^2$ -мерном пространстве, выделяемый следующей системой равенств и неравенств:

- (1) для всех i, j, p, q $X_{ijpq} = X_{jiqp}$,
- (2) для всех i $X_{i11} + X_{i22} = 1$,
- (3) для всех i $X_{i12} = 0$,
- (4) для всех i, j, p $X_{ijp1} + X_{ijp2} = X_{iipp}$,
- (5) для всех i, j, p, q X_{ijpq} неотрицательно,
- (6) для всех i, j, k, p, q, r сумма $X_{iipp} + X_{jkqr}$ больше либо равна сумме $X_{ijpq} + X_{ikpr}$.

Лемма 1. Для любого отображения f сегмента $\{1, 2, \dots, n\}$ в множество $\{1, 2\}$ точка X с координатами

$$\begin{aligned} X_{ijpq} &= 1, \text{ если } p = f(i) \text{ и } q = f(j), \\ X_{ijpq} &= 0, \text{ иначе,} \end{aligned}$$

является вершиной политопы P_n . Заметим, что такие точки соответствуют n -кликкам полного n -дольного графа с двумя вершинами в каждой доле.

Доказательство. Координаты точки X равны 0 или 1. При этом

$$X_{ijpq} = 1 \text{ тогда и только тогда, когда } X_{iipp} = 1 \text{ и } X_{jjqq} = 1.$$

$$\text{Если } i > j \text{ и } X_{ijpq} = 0, \text{ то } X_{iipp} = 0 \text{ или } X_{jjqq} = 0.$$

Равенства (1) и (3) очевидны. Равенства (2) следуют из того, что в каждой сумме одно слагаемое равно единице, а другое равно нулю.

Равенства (4) следуют из того, что в каждой сумме не более одного слагаемого равно единице, а другие равны нулю.

Проверим неравенство (6). Если $X_{ijpq} = X_{ikpr} = 1$, то обе части неравенства (6) равны 2. Если среди чисел X_{ijpq}, X_{ikpr} одно равно единице, то левая и правая части равны единице. Если координаты $X_{ijpq} = X_{ikpr} = 0$, то правая часть неравенства (6) равна нулю и не превосходит левую.

Итак точка X принадлежит политопу P_n . В точке X каждое из неравенств (5) обращается в равенство. Поэтому она является единственной точкой пересечения некоторых фасет политопы P_n . Следовательно, точка X является вершиной. Лемма доказана.

n -Дольному графу G , имеющему по две вершины в каждой доле, сопоставим аффинное подпространство $H(G)$, выделяемое всеми уравнениями $X_{ijpq} = 0$, где индекс i не равен индексу j и p -я вершина i -й доли не соединена ребром с q -й вершиной j -й доли.

Точке X политопы P_n сопоставим n -дольный граф $g(X)$, имеющий 2 вершины в каждой доле, у которого p -я вершина i -й доли соединена ребром с q -й вершиной j -й доли, где $j > i$, если $X_{ijpq} > 0$. Очевидно, если точка X принадлежит пересечению политопы P_n с подпространством $H(G)$, то граф $g(X)$ является подграфом графа G .

Удобно считать, что положительная координата точки X политопы P_n равна весу вершины или ребра графа $g(X)$.

Лемма 2. Если точка X принадлежит политопу P_n , то граф $g(X)$ содержит n -клик. Более того, эта клика может быть явно описана за время, ограниченное полиномом от n .

Доказательство. Выделим в каждой из долей графа $g(X)$ некоторую вершину, принадлежащую n -клик. При этом удобно рассматривать такую нумерацию вершин внутри доли, чтобы в клике оказались первые вершины каждой доли. Пусть для любого индекса i координата X_{i11} не меньше координаты X_{i22} .

Согласно (2), множество индексов долей $\{1, \dots, n\}$ равно объединению двух множеств

$$S_0 = \{i \mid X_{i11} > 1/2\}$$

$$S_1 = \{i \mid X_{i11} = X_{i22} = 1/2\}$$

Из равенств (4) следует, что для любых индекса i из множества S_0 и любого индекса j не равного i выполнено

$$X_{ij11} = X_{ii11} - X_{ij12} > 0.$$

Более того, если индекс j принадлежит множеству S_1 , то

$$X_{ij12} = X_{ii11} - X_{ij11} > 0.$$

Поэтому первая вершина в каждой доле из S_0 связана ребром с первой вершиной в каждой другой доле из S_0 и с каждой вершиной в любой доле из S_1 . Если множество S_1 содержит не более одного элемента, то искомым полным подграфом является первая вершина в каждой доле.

Если для индексов i и p число $X_{iipp} > 0$, то p -я вершина i -й доли в графе $g(X)$ соединена рёбрами с некоторой вершиной каждой другой доли в силу (4). Покажем, что в графе $g(X)$ есть такая клика, которой каждая доля из S_1 содержит вершину, принадлежащую этой клике. Если множество $S_1 = \{i, j\}$ содержит два элемента, то найдутся такие индексы p и q , что $X_{ijpq} > 0$.

Пусть множество S_1 содержит больше двух элементов. Из равенств (4) следует, что для любых индексов i, j из S_1 для любого индекса p существует такой индекс q , что X_{ijpq} не меньше $1/4$.

Фиксируем элемент a из множества S_1 . Можно считать, что для всех индексов j из $S_1 \setminus \{a\}$ X_{aj11} не меньше X_{aj12} . Множество $S_1 \setminus \{a\}$ является объединением двух непересекающихся множеств

$$S_{10} = \{j \text{ из } S_1 \setminus \{a\} \mid X_{aj11} > 1/4\}$$

$$S_{11} = \{j \text{ из } S_1 \setminus \{a\} \mid X_{aj11} = X_{aj12} = 1/4\}$$

Из неравенств (6) следует, что для всех индексов i, j, k из S_1 и для всех индексов p, q, r сумма $X_{ijpq} + X_{ikpr}$ не превосходит сумму $1/2 + X_{jkqr}$.

Для любых разных индексов i и j из S_{10} первые вершины в a -й, i -й и j -й долях соединены рёбрами. Более того, для любых индексов i из S_{10} и j из $S_1 \setminus \{a, i\}$ первые вершины в a -й, i -й долях соединены рёбрами с каждой вершиной j -й доли. Если множество S_{11} пусто или содержит единственный элемент, то искомая клика включает первые вершины каждой доли. Поэтому остаётся построить клику, вершины которой лежат в долях с номерами из множества S_{11} , когда оно имеет не меньше двух элементов. Для этого фиксируем элемент b из S_{11} и разбиваем множество $S_{11} \setminus \{b\}$ в объединение двух непересекающихся подмножеств S_{110} и S_{111} . Повторяя этот процесс, мы либо построим n -клик в графе $g(X)$, либо за не более чем n шагов придём к задаче построения клики, вершины которой лежат в долях с номерами из такого множества S , что для любых индексов i, j из S и любых индексов p, q координата $X_{ijpq} = 1/4$. Очевидно, что можно выбрать любую вершину в каждой доле из S . Лемма доказана.

Заключение

Теорема. Пусть n -дольный граф G имеет по две вершины в каждой доле.

Если пересечение I политопы P_n и пространства $H(G)$ не пустое, то граф G имеет хотя бы одну n -клик.

Если размерность пересечения I политопа P_n и пространства $H(G)$ равна единице, то граф G имеет либо одну, либо две n -клики.

Доказательство. Из лемм 1 и 2 следует, что граф G содержит n -клику тогда и только тогда, когда пересечение политопа P_n и пространства $H(G)$ непустое. Более того, если найдена точка X в пересечении политопа P_n с пространством $H(G)$, то n -клика определяется эффективно по лемме 2. Политоп P_n задан системой из $O(n^3)$ равенств и неравенств. Поиск n -клики сводится к задаче линейного программирования и требует лишь полиномиального времени.

С другой стороны, n -клики соответствуют вершинам политопа P_n . Если граф G имеет и n -клики, то пересечение I содержит три вершины политопа. Поскольку никакая из вершин не является выпуклой комбинацией двух других, то размерность пересечения I не меньше 2. Противоречие доказывает, что существует не больше двух n -клик.

Отметим, что размерность политопа P_n равна $n(n+1)/2$. Действительно, поскольку координаты вида X_{ijl} однозначно определяют точку политопа P_n , то размерность не превосходит $n(n+1)/2$.

С другой стороны, среди вершин, соответствующих n -кликам полного n -дольного графа, существует $n(n+1)/2+1$ аффинно независимая точка. Это точка с координатами $X_{ijl} = 0$, n точек, у которых среди координат X_{ijl} одна единица, а остальные равны нулю, и $n(n-1)/2$ точек, у которых среди координат X_{ijl} две единицы, а остальные равны нулю. Поэтому размерность политопа P_n равна $n(n+1)/2$.

Список литературы

- [1] Even S., Itai A., Shamir A. On the complexity of timetable and multicommodity flow problems. // SIAM Journal on Computing, v. 5, n. 4, 1976, p.p. 691-703.
- [2] Схрейвер А. Теория линейного и целочисленного программирования. М.: Мир, 1991.

АЛГОРИТМ ПОИСКА СИГНАЛА СО СЛОЖНОЙ СТРУКТУРОЙ

Е.В. Любецкая¹, Л.А. Леонтьев¹, М.А. Ширшин², В.А. Любецкий¹

¹Институт проблем передачи информации, Российская академия наук
127994, Москва, Большой Каретный переулок, 19
lin@iitp.ru, leontiev@iitp.ru, lyubetsk@iitp.ru

²Государственная Классическая Академия имени Маймонида
115035, Москва, Садовническая улица, 52/54
shirshin@tblabs.com

Ключевые слова: *аттенюация, регуляция транскрипции, выделение сигнала.*

Abstract

The task of regulator signal detecting is formulated for a signal with complicated structure. The algorithm solving this task is presented.

Введение

Текущие концентрации молекул в клетке в значительной мере зависят от протекающих биохимических реакций в ней (в работе рассматривается случай прокариот). Как правило, реакция протекает по мере поступления в цитоплазму клетки соответствующего набора ферментов, что зависит от экспрессии соответствующих групп генов (регулонов и их полицистронных случаев – оперонов). Таким образом, текущая жизнь клетки в значительной степени состоит в регуляции групп генов в зависимости от ее внутренних и внешних условий жизнедеятельности. Известны разные типы регуляции: основанные на белок-ДНКовом взаимодействии (позитивная или негативная, когда активатор или репрессор связывается с соответствующим сайтом в лидерной области оперона, обычно расположенным внутри промотера или вблизи него); или основанные на образовании специфических вторичных структур мРНК (например, альтернативных и, в частности, аттенюаторных – в последнем случае механизм регуляции зависит от взаимного расположения РНК-полимеразы и рибосомы в параллельно идущих процессах транскрипции и трансляции); регуляция аллостерическая (когда конечный продукт катализируемой ферментами реакции ингибирует работу одного из ферментов) и другие. В процессах мРНК-овой регуляции большую роль играют: лидерный пептид с регуляторными кодонами, стабилизирующие белки, тРНК, молекулы-эфффекторы и т.п. Часто в управлении одного оперона участвуют несколько разных типов регуляции.

1 Постановка задачи

Регуляции с помощью белков-репрессоров или активаторов, а также аллостерические регуляции изучаются сравнительно давно. Фундаментальная важность альтернативной регуляции обнаружена недавно, когда были найдены новые ее примеры.

В настоящее время расшифровано и доступно более 100 полных геномов, несколько сотен полных геномов секвенируются и будут доступны в ближайшее время, не говоря уже о секвенировании частей геномов. Такой огромный объем информации делает невозможным лабораторный биохимический анализ подавляющего большинства геномов, поэтому необходимы **алгоритмы** компьютерного анализа геномов и, в частности, *поиска потенциальных аттенюаторных структур* в достаточно полно секвенированных геномах, которые были бы применимы для **массового анализа** сразу всех организмов из данной таксономической группы.

В основном для поиска *регуляторных сигналов* до сих пор применялись **два подхода**: составлялось распознающее правило (по выборке лидерных областей, содержащих достаточно сходные регуляторные сайты); или такой сигнал искался непосредственно в каждой из последовательностей, входящих в выборку, на основе существенной консервативности сигнала. Оба эти подхода плохо применимы в случае массового поиска аттенюаторных структур. Ситуация особенно усложняется, когда речь идет о поиске регуляторного сигнала для генов с неизвестной функцией или для геномов, у которых еще не выяснена структура интересующего нас оперона.

Нами рассматривается задача поиска аттенюаторной регуляции в лидерной области гена. Математическая постановка этой задачи такова. Дана последовательность фиксированной длины в алфавите $\{A, C, T, G\}$; нужно найти определенную уточняемую ниже комбинацию подслов в ней, соответствующую биологическому понятию аттенюаторной структуры. А именно, аттенюаторная структура искалась как состоящая из лидерного пептида с регуляторными кодонами, участка остатков урацила и трех шпилек: терминатора, антитерминатора и паузной, определенным образом расположенных относительно друг друга. К этому традиционному описанию нами было добавлено представление о трех словах – «основах» этих трех шпилек. Наш алгоритм реализует идею поиска именно таких трех слов, «порождающих» три соответствующие шпильки; описание этого приведено ниже.

2 Описание алгоритмов

Условимся *обозначать* плечи паузной шпильки 1 и 2, плечи терминатора 3 и 4, а плечи антитерминатора 2' и 3'.

Алгоритм применяется для исследования вопроса о наличии аттенюаторной регуляции в **одной** данной лидерной области: в ней ищется сама аттенюаторная структура или обоснование ее отсутствия (т.е. ищется алгоритмическое указание на отсутствующий элемент такой структуры). Нами обычно рассматривались случаи, когда лидерная область имеет длину в 450-800, а иногда и 1000 нуклеотидов. Алгоритм ищет *структуру*, состоящую из следующих элементов в исходной нуклеотидной последовательности: рамки считывания, кодирующей лидерный пептид с набором регуляторных кодонов в нем; тройки слов s_1, s_2, s_3 , которые служат «основой» для построения плеч: s_1 плеч 2 и 2', s_2 плеч 3 и 3', s_3 плеча 4; и трех шпилек – терминатора, антитерминатора и паузной (в соответствии с этими плечами); участок остатков урацила.

В алгоритме используются следующие основные параметры: минимальная длина лидерного пептида; длина участка, содержащего остатки урацила; минимальное количество самих остатков урацила в этом участке; длина каждого из слов s_1, s_2, s_3 ; максимальное расстояние (разница между началом первой буквы следующего слова и последней буквой предыдущего слова) между словами s_1 и s_2, s_2 и s_3 ; максимальное число различных (в паре s_1 и s_3) и некомплементарных (в парах s_1 и s_2, s_2 и s_3) нуклеотидов; максимальное расстояние между началом слова s_3 и концом слова s_2 ; максимальное число отбираемых алгоритмом троек слов s_1, s_2, s_3 ; максимальное расстояние между началом участка остатков урацила и концом слова s_3 ; минимальное отношение числа пар GC к числу пар AT в словах s_2 и s_3 ; интервальное значение длины фрагмента до начала петли антитерминатора – паузная шпилька ищется в этом фрагменте.

Значения этих параметров подбирались с помощью стандартных приемов статистики, один из них (поиск доверительного интервала с помощью распределения Стьюдента) в качестве примера указан ниже. Для наших результатов характерны следующие значения этих параметров. Они перечисляются в том же порядке, как выше: 30 нуклеотидов, 5, 8, 100, 2, 30, 15, 9, 1.3, от 100 до 50.

Алгоритм последовательно находит.

1. Кандидатов в лидерные пептиды. Для этого перебираются все открытые рамки считывания, в каждой из которых ищется скопление регуляторных кодонов соответствующей аминокислоты.

кислоты (в процессе счета список этих кодонов может расширяться – он также является параметром алгоритма).

Открытой рамкой считывания считается участок последовательности, удовлетворяющий условиям:

1. Его первый кодон является старт-кодоном (обычно это ATG, реже GTG или TTG);
2. Последний кодон является стоп-кодоном (одним из кодонов TAA, TAG, TGA).
3. Длина участка в нуклеотидах кратна 3 (т.е. старт-кодон, стоп-кодон и кодоны, лежащие между ними, находятся в рамке считывания).
4. Расстояние между старт и стоп-кодонами не менее *мин_дл_пепт*. Поиск кандидатов выполняется тривиально:

А. В последовательности находим стоп-кодона. Отступая влево от каждого стоп-кодона на расстоянии, большем *мин_дл_пепт* и кратном 3, ищем старт-кодона. Исключаем рамки считывания содержащие другие стоп-кодона.

В. Стоп-кодона, для которых не найден старт-кодон, отсеиваются. Для каждого стоп-кодона находится множество всех возможных старт-кодонов, из которого выбирается один с наибольшей координатой начала.

С. Для каждой пары старт-кодон и стоп-кодон находящиеся между ними триплеты размечаются как кодоны соответствующих аминокислот.

Д. Среди так полученных рамок считывания отбираем кандидатов в лидерные пептиды. Для этого в каждом рамке ищем непрерывное поле соответствующих регуляторных кодонов. Если специфической для данного оперона аминокислоте соответствует малое количество различных кодонов (например, 1-2), то поле может быть более коротким (2 и более кодона). Если аминокислоте соответствует более 2-х кодонов, то ищем поле из 4 и более кодонов. Известно, что некоторые опероны регулируются с помощью лидерных пептидов, содержащих в таком поле специфические кодоны «чужих» аминокислот (например, биосинтез треонина является изолецин-зависимым и т.д.) – в таком случае появление в поле кодонов «чужих» аминокислот не считается разрывом.

Если ни в одной рамке не нашлось поля специфических кодонов, алгоритм выдает сообщение об отсутствии аттенуаторной регуляции.

Для каждого найденного кандидата в лидерные пептиды ищутся все подходящие участки остатков урацила.

Для каждой пары таких объектов ищутся тройки слов s_1, s_2, s_3 ; тройки сортируются в порядке увеличения расстояния между s_2 и s_3 , из них отбирается заданное число первых. Тройку слов $\langle s_1, s_2, s_3 \rangle$ ищем как комбинацию подслов последовательности одинаковой длины *длина_слова* (обычно *длина_слова* = 8), находящихся на расстоянии друг от друга не более константы *макс_длина_шпильки* (обычно *макс_длина_шпильки* = 100) нуклеотидов и удовлетворяющих условиям:

1. Слова s_1, s_2 комплементарны с точностью до замены любых *макс_разница* (*макс_разница* < *длина_слова*) нуклеотидов.
2. Слова s_2, s_3 комплементарны с точностью до замены любых *макс_разница* (*макс_разница* < *длина_слова*) нуклеотидов.
3. Слова s_1, s_3 совпадают с точностью до замены любых *макс_разница* (*макс_разница* < *длина_слова*) нуклеотидов.

Для каждого подслова s_1 длины *длина_слова* на отрезке $[s_1 + \text{длина_слова}; s_1 + \text{длина_слова} + \text{макс_длина_шпильки}]$ ищутся комплементарные слова s_2 . Если находятся, то для каждого слова s_2 на отрезке $[s_2 + \text{длина_слова}; s_2 + \text{длина_слова} + \text{макс_длина_шпильки}]$ ищутся комплементарные слова s_3 . Далее проверяется совпадение слов s_1 и s_3 ; выдаются тройки.

Из всех полученных троек слов выбираем одну, выполняя следующие действия.

Отсортируем эти тройки по возрастанию разницы между координатами первой буквы слова s_3 и последней буквы слова s_2 . Ищем слова, удовлетворяющие следующим условиям:

1. Разница между координатами первой буквы слова s_3 и последней буквы слова s_2 меньше *макс_дельта* (обычно *макс_дельта* = 30; иначе не образуется шпилька-терминатор), а среди них отбираются первых *количество_ответов* ответов (обычно *количество_ответов* = 15)

2. Координата начала слова s_3 меньше координаты начала участка урацилов U и отстоит от нее не более, чем на *расстояние_до_поля_T* нуклеотидов (обычно *расстояние_до_поля_T* равнялось 9).

Алгоритм отслеживает и такое дополнительное условие:

3. Суммарное число комплементарных пар GC в паре слов s_2 и s_3 , поделенное на число комплементарных пар AT не меньше x (обычно $x = 1.3$), т.е. эта пара слов должна быть GC-насыщенной.

В случае нахождения такой тройки алгоритм переходит к следующему шагу и сообщает «аттенуаторная регуляция обнаружена». Если такая тройка не находится, то алгоритм выдает сообщение об отсутствии аттенуации.

Затем последовательно по s_2 и s_3 строится терминатор из условия минимума энергии, и аналогично – антитерминатор и паузная шпилька (при заданных параметрах).

Среди так полученных структур отбираются по одному представителю из каждого класса «подобных» структур (определяемого нами отношением эквивалентности). Для наших случаев, как правило, фактически оказывалось по одному-двум, но не более трех таких представителей.

Итак, терминатор образуется спариванием слов s_2 и s_3 и продолжением спаривания в обе стороны. Альтернативность терминатора и антитерминатора поддерживается за счет слова s_2 : антитерминатор – это шпилька, включающая спаривание s_1 и s_2 с изменением свободной энергии сворачивания, меньшим некоторого параметра (например, -10 ккал/моль, что обеспечивает возможность сворачивания антитерминатора).

3 Обсуждение алгоритма и обзор результатов

Правильность работы алгоритма подтверждается сравнением результатов счета с известными биологическими данными, а также результатами совместных выравниваний найденных нами новых и уже известных сигналов аттенуаторной регуляции. Выравнивания говорят, в частности, о консервативности указанных троек слов. Действительно:

1. При нашем выравнивании для многих оперонов и независимо от алгоритма определяются слова s_1 , s_2 , s_3 и они оказываются консервативными. Так, для *trp*-оперона эти слова выравнивались даже у гамма-, альфа- и бета-протеобактерий. Для каждого из 8 изученных оперонов у организмов из одной группы эти слова оказывались одинаковыми с точностью до нескольких букв.

2. Для многих оперонов (например, *thrABC* у гамма-протеобактерий) антитерминаторы, полученные алгоритмом для разных организмов, имели сходную структуру: близкие значения числа отрезков и их длин, близкие значения длин выпячиваний и их типов (односторонние или двусторонние), и т.д.

Алгоритм обеспечивает следующие условия:

1. Наличие лидерного пептида с полем регуляторных кодонов соответствующих аминокислот.

2. Пересечение терминатора и антитерминатора.

3. Отсутствие пересечения паузной шпильки и терминатора.

4. Расстояние от конца поля регуляторных кодонов до начала левого плеча антитерминатора более 5 нуклеотидов.

Кроме того, алгоритм проверяет полученные им ответы на выполнение условия, которое наблюдалось во многих известных случаях аттенуаторной регуляции: расстояние от конца лидерного пептида до начала левого плеча антитерминатора от -3 до $+3$ нуклеотидов.

Упомянутое хорошее выравнивание по разным причинам не является обязательным условием наличия аттенуаторной регуляции. В некоторых организмах (например, *Y. pestis*) фрагменты последовательности, участвующие в такой регуляции, не являются консервативными в других гамма-протеобактериях.

Для нахождения численных значений перечисленных выше параметров алгоритма применяются стандартные методы статистики. По известным значениям X этих параметров, которые, в основном, брались из результатов экспериментальных работ, а также из наших уже подтвержденных выравниванием результатов, находились значения параметра и доверительный интервал. Распределение X было близко к нормальному, что проверялось стандартным образом; или в некоторых случаях мы работали с другими (например, log-нормальным) распределениями. За само значение бралось среднее X^* (обычно, среднее арифметическое, иногда иное среднее) и находилось квадратичное отклонение s . Обычно выбиралась надежность $\gamma = 0,95$ и искался

радиус доверительного интервала, равный $t_\gamma \cdot \frac{s}{\sqrt{n}}$, где n – объем наблюдаемых значений и t_γ находился стандартным образом по таблице, отражающей соотношения

$$P\left(\left|\frac{X^* - a}{s/\sqrt{n}}\right| < t_\gamma\right) = 2 \int_0^{t_\gamma} S(t, n) dt = \gamma. \text{ Здесь } S(t, n) \text{ - плотность распределения Стьюдента.}$$

Для наших параметров значения радиуса было небольшим.

БИОИНФОРМАТИКА: ОТ ОБСЛУЖИВАНИЯ ЭКСПЕРИМЕНТА К ПОЛУЧЕНИЮ БИОЛОГИЧЕСКИ ЗНАЧИМЫХ РЕЗУЛЬТАТОВ

А.А. Миронов

Московский государственный университет
119992, Москва, Воробьевы горы
aa_mironov@pochtamt.ru
тел: +7 (095)939-43-31

Ключевые слова: биоинформатика, анализ последовательностей, анализ данных

Abstract

This paper describes history of bioinformatics and presents some biologically significant results that obtained purely by computer analysis.

Введение

В настоящее время биоинформатика, как следует из названия, является наукой, находящейся на стыке биологии и информатики. Основные задачи, которыми занимается эта наука - анализ последовательностей.

1 История биоинформатики

Биоинформатика оформилась как самостоятельная дисциплина в начале 80-х годов XX века. Поводом к возникновению этой области науки послужило открытие эффективного метода определения нуклеотидных последовательностей. В 1978 - 80 гг. были определены несколько полных последовательностей размером около 4 тыс. оснований (букв) – плазмиды pBR322, вируса саркомы зеленой марьшанки SV40, бактериофага ФХ-174. Стало ясно, что последовательности длиной в несколько тыс. оснований не могут быть проанализированы вручную. И появилась необходимость в компьютерных методах обработки таких массивов информации. Задачи, которые тогда стояли часто были очень простыми – оттранслировать нуклеотидную последовательность в белок, найти открытую рамку, найти сайты рестрикции, найти сходство двух последовательностей. Этот круг задач определялся потребностями экспериментаторов. Хотя уже в то время стали формулироваться и более сложные задачи, некоторые из них до сих пор нельзя считать окончательно решенными, – поиск генов, физическое картирование, поиск сигналов в последовательностях, предсказание вторичной структуры РНК. Биоинформатика оформилась как самостоятельная, хоть и вспомогательная, ветвь науки в январе 1980 г., когда вышел специальный компьютерный выпуск биологического журнала “Nucleic acids research”.

Впрочем, первые биоинформатические работы, посвященный анализу последовательностей, появились раньше (в конце 60-х - начале 70-х годов XX века). Это теперь уже классические работы по сравнению аминокислотных последовательностей белков - построение матрицы сравнения аминокислотных остатков (Дейхофф) и построение алгоритма выравнивания последовательностей (Нидельман, Вунш). Другое направление исследований было связано с анализом возможных вторичных структур транспортных РНК. В результате Эйген обнаружил, что все тРНК можно уложить в характерную структуру, похожую на клеверный лист. Этот результат сразу вошел в учебники по молекулярной биологии, и только через 12 лет было экспериментально показано, что это предсказание верно.

Следующий этап связан с увеличением количества прочитанной информации. Количество расшифрованных последовательностей перевалило за тысячу, размеры последовательностей

стали достигать сотни тысяч. Это привело к тому, что возникла необходимость хранения и доступа к информации. Так появились первые банки последовательностей - банк Лос-Аламос, база данных EMBL, базы данных по аминокислотным последовательностям. Появился новый класс задач - быстрый поиск сходства в банках последовательностей.

Следует отметить, что скорость роста количества прочитанной информации примерно совпадала со скоростью роста мощности компьютеров - удвоение за полтора года. Однако некоторые задачи имеют нелинейную зависимость сложности от размера данных. Поэтому по-прежнему остро стоят вопросы создания эффективных алгоритмов.

Ситуация стала меняться с появлением проектов расшифровки полных геномов, в том числе генома человека. Если до этого периода задачи ставились так - есть последовательность, про которую уже многое известно. Найти что-нибудь еще. Теперь же компьютерный анализ выходил на первый план. Требовалось проанализировать полный геном и найти в нем новые особенности - новые гены, сигналы регуляции и т.п. И если раньше компьютерный анализ шел после эксперимента, то теперь результаты компьютерного анализа часто определяют направление экспериментальной работы. Более того, часто результаты компьютерного анализа являются самостоятельными биологически значимыми результатами.

Существенную роль в получении биологически значимых результатов сыграло появление не одного, а нескольких родственных геномов. Это позволяет делать сравнительный анализ геномов, который, в свою очередь, является очень мощным подтверждающим или отсекающим контролем.

В биоинформатике весьма важным этапом является формулировка задачи. Если на биологическом уровне задача может быть (не всегда) достаточно хорошо поставлена, то математическая формулировка допускает большую степень неоднозначности. Типичный пример - требуется определить похожи ли две последовательности? Простейшая математическая формулировка соответствующей задачи - найти редакционное расстояние. Однако, корректное решение этой математической задачи часто приводит к неправильным биологическим результатам. Следовательно, необходимо так сформулировать математическую задачу, чтобы ее решение приводило к адекватным результатам. Здесь математика также диктует свои условия. Эта формулировка должна позволять эффективно решать соответствующую математическую задачу.

2 Примеры применения биоинформатики

Пример удачного применения компьютерного анализа. Анализ альтернативного сплайсинга в геноме человека. Краткое биологическое введение. Многие гены человека и других организмов (за исключением бактерий) имеют мозаичную структуру. В них значащие (кодирующие) сегменты (экзоны) прерваны незначащими (некодирующими) сегментами (интронами). Причем часто размеры интронов значительно превышают размеры экзонов. В клетке существует специальный механизм, который узнает интроны и их вырезает. Он называется сплайсингом. Ранее были известны отдельные природные случаи, когда из одного гена собиралась разная мозаика экзонов, что приводило к образованию разных белковых продуктов. Однако долгое время считалось, что это редкие экзотические случаи.

Существующие компьютерные методы могут по последовательности гена узнать интроны, и предсказать мозаику экзонов. Однако они ненадежны и работают гораздо менее точно, чем молекулярная машина клетки. Но в рамках проекта Геном человека, существовала программа определения последовательности не только всего генома, но также и уже сплайсированных РНК, а точнее их фрагментов - EST. Надо отметить, что определение последовательности генома, или EST сейчас носит промышленный характер - существуют автоматизированные фабрики по чтению геномов. Количество определенных EST исчисляется миллионами.

Компьютерная задача, которую мы решали, заключалась в том, чтобы положить множество EST на геном человека. Поскольку EST сплайсированы (не содержат интронов), а геном -

нет, то такое картирование позволяет определить в геноме интроны. Впрочем, когда мы эту задачу делали, генома человека еще не было, а были определены последовательности довольно большого количества отдельных генов. Был разработан эффективный алгоритм картирования фрагментов РНК (EST) на геном, и к нашему удивлению мы обнаружили большое количество неоднозначностей, когда одни EST указывают на то, что определенный фрагмент гена является интроном, а другие говорят, что это экзон. Поскольку данные EST достаточно грязные (по построению), то мы ввели значительное количество фильтров, уменьшающих влияние артефактов.

В результате обнаружилось, что до половины генов (вместо 6% по прежним оценкам) имеют разные варианты сборки мозаики экзонов. Когда мы это обнаружили, то были проблемы с публикацией результатов, однако теперь это общее место. Все знают, что как минимум половина генов подвержены альтернативному сплайсингу. Биологически этот феномен можно интерпретировать по-разному. С одной стороны, альтернативный сплайсинг позволяет Природе конструировать новые специальные формы белков из имеющихся кирпичиков. С другой - заметная доля изоформ не кодирует белков, и поэтому может быть просто ошибкой клеточной машины.

Далее мы решили сравнить альтернативный сплайсинг человека и мыши. С точки зрения молекулярной биологии это весьма близкие организмы. 99% генов одного организма представлен в другом. Большинство генов сплайсируется одинаково. Мышь часто используется в качестве модельного организма для исследования человека. В частности, многие исследования рака, предварительные испытания лекарственных препаратов и многое другое делается на мышах.

Изоформы, существующие у человека, были картированы на геном мыши. Если при этом оказывалось, что изоформа не может быть картирована полностью, а некоторые экзоны пропадали, то делался вывод о том, что соответствующего экзона у мыши нет, и мы имеем дело с неконсервативным альтернативным сплайсингом. В результате анализа оказалось, что более 30% генов человека имеют видоспецифические изоформы. Встречный анализ показал, что около 30% генов мыши имеют свои специфические изоформы.

Существует еще множество примеров применения компьютерных методов анализа, приведших к значимым результатам - это и поиск новых генов, и анализ регуляции, и поиск новых регуляторных сигналов, и многое другое.

Резюмируя, можно сказать, что в настоящее время биоинформатика стала существенным элементом биологической науки, способной не только обслуживать потребности экспериментаторов, но также и продуцировать собственные результаты.

ИНФОРМАЦИОННОЕ УПРАВЛЕНИЕ ПРОЦЕССОМ ИЗУЧЕНИЯ СЛОЖНОЙ РЕАКЦИИ

С.Н. Истомина

Авиационно-технологический университет им Циолковского
121609, Москва, Оршанская ул., 8а
lyubetsk@iitp.ru
тел: +7 (095) 413-46-43, факс +7 (095) 209-05-79

Ключевые слова: биоинформатика, химические реакции дисперсионный анализ, регрессионный анализ

Abstract

We propose a new approach to study complex reactions based on step-by-step informational management for a research process of some object.

Введение

Многие задачи в управлении, науке и технике (далее примеры берутся из химической технологии) приводят к необходимости исследовать процесс, в котором параметр y меняется в зависимости от нескольких факторов x_1, x_2, \dots, x_k . Желательная цель такого исследования состоит в том, чтобы получить математическое описание этой зависимости, в частности, для последующего управления этим процессом.

Известно, что протекание сложной органической реакции основано на сигнальном взаимодействии, выражающемся вторичной структурой молекул. В реакции участвуют реагенты, катализаторы и растворители с уже известной структурой, что позволяет записать не только стехиометрическое уравнение реакции, но и рассмотреть несколько альтернативных схем протекания реакций с соответствующими кинетическими закономерностями.

Дисперсионный анализ позволяет по эмпирическим сгруппированным данным выделить те факторы, которые значимо влияют на изучаемый параметр. При достаточном количестве данных это позволяет указать наличие взаимодействия между парами, тройками и более факторов и ранжировать все факторы и взаимодействия. Существенным недостатком является необходимость большого количества эмпирических данных, но и при их наличии дисперсионный анализ не предлагает описания самой зависимости.

Регрессионный анализ для описания связи между интересующими нас величинами предлагает построить многочлен, который является спрямлением эмпирических данных. При построении уравнения регрессии минимизируется сумма квадратов отклонений эмпирических данных от получаемой регрессии.

Представляется плодотворным объединить два упомянутых подхода в схеме, которая позволит целенаправленно строить регрессию и при этом поэтапно и обоснованно расширять знания о процессе, который описывается этой регрессией. При этом актуальной становится задача согласования эмпирических данных, полученных на разных этапах, для получения адекватного описания их уравнением регрессии.

Классический подход для изучения реакции заключается в изучении изменения скорости реакции или выхода интересующего продукта при поочередном изменении одного из условий протекания этой реакции. Оптимальным является варьирование многих факторов на определенном числе уровней.

Поэтому неизбежным становится применение оптимального планирования эксперимента с отвечающими поставленной задаче критериями оптимальности (например, ортогональность планов, то есть условий проведения опытов).

Такой подход можно назвать информационным управлением самим процессом изучения некоторого явления, процессом построения математической модели, например сложной химической реакции. Он позволяет использовать информацию, собранную разными способами обработки данных в обширной факторной области и получить выводы, согласующиеся со всем комплексом знаний о процессе.

1 Модель

Построение регрессии принято начинать с линейного уравнения и, если оно недостаточно спрямляет данные, то последовательно увеличивать число членов многочлена, включая произведения факторов, их степени и т.д. Вместо этого предлагается последовательно включать члены уравнения, значимость которых для эмпирических данных определяется дисперсионным анализом или теоретическими положениями. При этом подходе в уравнение регрессии часто не будут включаться коэффициенты при высших степенях. Получаемые при этом уравнения регрессии имеют преимущества при анализе и интерпретации зависимости между изучаемыми величинами. Их особенностью является то, что при фиксированных значениях всех факторов, кроме одного, зависимость по этому оставшемуся переменным фактору является линейной (а уравнение часто называют полилинейным). Такие модели пригодны для описания стационарной области, позволяют анализировать влияние одних факторов при изменении уровней других, то есть получить количественную информацию о взаимодействиях факторов, и в то же время их построение требует меньшего числа эмпирических данных (например, можно использовать двухуровневые ортогональные планы).

При переходе к изучению новой области протекания процесса указанный подход позволяет использовать имеющиеся эмпирические данные и уже построенную модель в качестве исходной. На следующем этапе целенаправленно, с учетом полученной информации, планируют и проводят опыты, на основании которых получают уточненное описание исследуемой зависимости, после чего переходят к следующему этапу. Хотя и весьма частными примерами применения предлагаемой схемы исследования могут служить:

Изучение синтеза нового технологически важного продукта 2-метиленициклогексанола из циклогексанона.

2 Результаты

Изучение конкурентной реакции синтеза N-винилацетанилида, с параллельной реакцией приводящей к побочному, но промышленно перспективному продукту 3-фениламино-1-бутина.

1. Исследовались условия протекания реакции восстановительного метилирования кетон системой КОН-ДМСО проводилось с целью изучения условий синтеза с тем, чтобы повысить выход при технологически приемлемых условиях, а также получить представления о типе реакции.

Описанным выше методом для получения математической модели процесса выведено уравнение регрессии, адекватно описывающее эмпирические данные. Анализ уравнения позволил указать условия наибольшего выхода основного продукта. Затем были спланированы и осуществлены уточняющие опыты, на основании которых сделан вывод о том, что достигнута стационарная область протекания реакции со стабильно хорошим процентным выходом N-винилацетанилида. Дальнейшее исследование проводилось для более детального изучения протекания реакции, чтобы получить представление о кинетике и механизме реакции в найденной нами стационарной области. На основе полученной информации появилась возможность более точно описать стационарную область и уточнить уравнение регрессии, а также, по

возможности, сделать новые выводы. Все множество опытов составляло реплику от четырехфакторного плана с 3 уровнями по трем факторам, и с 4 уровнями по четвертому (времени протекания реакции) – для последующих расчетов скоростей реакции. Отсюда было получены кинетические представления о ходе реакции. Получено полиномиальное уравнение третьего порядка, не содержащее высших степеней, но со значимым коэффициентом при тройном взаимодействии. Гипотеза об адекватности не отвергается.

Уравнение регрессии были использовано для получения оценок скоростей реакции в нескольких точках факторного пространства, которые затем были использованы для получения предварительных оценок кинетических констант и порядка реакции. Согласие статистической модели изучаемого процесса и качественного кинетического описания механизма реакции подтверждает последние выводы.

2. На основе того же метода проводилось изучение реакции ацетилен с ацетанилидом с целью найти условия раздельного протекания реакции по двум направлениям. Поэтому здесь рассматривались два параметра – выходы N-винилацетанилида и 3-фениламино-1-бутина соответственно. Были получены два уравнение регрессии, описывающие зависимость выхода основного и побочного продукта отдельно, но при одинаковых условиях. При этом были использованы полиномиальные уравнения, что позволило найти области раздельного протекания реакции в пространстве трех факторов. Анализ уравнения позволил указать условия наибольшего выхода основного продукта практически лишь со следами образования побочного продукта. Также были найдены условия протекания реакции, главным образом, по ее второму направлению. При этом удалось повысить выход реакции по второму продукту при небольшом процентном образовании основного. Затем были спланированы и осуществлены опыты, на основании которых стало возможным оценить скорости протекания реакции в раздельных областях по рассматриваемым параметрам и сделаны предварительные оценки кинетических констант и выводы о механизме конкурентных реакций. В этом примере, как и в предыдущем, согласие статистических моделей изучаемого процесса и качественного кинетического описания механизма реакции подтверждает выводы.

Широко известна задача моделирования биосинтеза веществ в бактериальной культуре и, в частности, построения модели регуляции экспрессии генов, см., например, [1]. Попытки такого моделирования, основанные на аппаратах дифференциальных уравнений или случайных процессов, сталкиваются с известными вычислительными трудностями, а также с трудностями интерпретации результатов. Последние связаны с косвенным характером описания биологических процессов при использовании такого математического аппарата. Желательно развитие более прямого и вычислительно более прозрачного аппарата для решения этой задачи.

Мы предлагаем подход к указанной задаче, основанный на обобщенной теории автоматов (ориентированный граф с потенциально бесконечным числом меток). На этом языке описываются процессы транскрипции и трансляции – в этом сообщении в части репрессорной и аттенуаторной регуляций экспрессии оперона и биосинтеза аминокислот. А именно, предлагается выделить конечное число вершин графа – «состояний системы»: из одного состояния возможен переход в одно из нескольких других в зависимости от, например, концентрации веществ и значений некоторых случайных величин. Скажем, в зависимости от концентрации триптофана из «паузно-кодонного» состояния (когда РНК-полимераза находится в последнем такте паузы, а рибосома – на последнем регуляторном кодоне) возможен переход в «послепаузно-кодонное» состояние (когда рибосома осталась на том же регуляторном кодоне) или переход в «послепаузно-послекодонное» состояние (когда рибосома сошла с регуляторных кодонов). При этом вершинам графа на каждом такте динамически приписываются числа, характеризующие: долю бактерий из культуры, находящихся в данном состоянии, и некоторых средние показатели культуры. На основе такого подхода нами, в частности, компьютерно моделировалась регуляция биосинтеза триптофана. Были получены графики, которые показывают, например, что ре-

прессорная регуляция быстрее реагирует на концентрацию триптофана, чем аттенуаторная (факт экспериментально известный).

Список литературы

- [1] Истомина С.Н., Рубанов Л.И. Параллельный алгоритм поиска регуляторного сигнала в геномах бактерий // Электронный научный журнал «Информационные процессы» (<http://www.jip.ru>), т. 2, № 1, 2002, с.с. 85-90.

ИНСТРУМЕНТАЛЬНАЯ СИСТЕМА ВИЗУАЛЬНОГО ПРОЕКТИРОВАНИЯ «ONTOEDITOR»

О.А.Невзорова¹, В.Н. Невзоров²

¹НИИ математики и механики им. Н.Г. Чеботарева
420008, Казань, ул. Кремлевская, 18
E-mail: Olga.Nevzorova@ksu.ru
тел. +7 (8432) 92-62-69

²Казанский государственный технический университет
420111, Казань, ул. К. Маркса, 10
E-mail: nevzorov@mi.ru
тел.+7 (8432)31-00-81

Ключевые слова: онтология, визуальное представление, инструментальные средства

Abstract

In this paper we described the current version of the AviaOntology and the architecture of the development support system "OntoEditor". The main functions of this system and the methods of visual design are considered.

Введение

Информационный ресурс АвиаОнтология является лингвистической онтологией, разработанной для предметной области, описывающей поведение оператора (летного экипажа) и бортовой аппаратуры в различных полетных режимах, и представляет собой иерархическую сеть понятий предметной области. К настоящему времени АвиаОнтология содержит 1600 понятий, 4700 текстовых единиц. Проектирование подобной модели является сложной экспертной задачей, требует взвешенного подхода к выбору формализма описания.

В качестве базовой технологии для проектирования прикладной онтологии была выбрана разработанная в НИВЦ МГУ технология проектирования больших и сверхбольших онтологий и тезаурусов для различных предметных областей [1].

Основными текстовыми источниками знаний о рассматриваемой предметной области являются серия специальных книг, отобранных экспертами в данной области, и текстовая информация из Интернет по отдельным разделам предметных знаний. Общий объем текстовой коллекции в настоящий момент составляет свыше 100 Мб.

АвиаОнтология предназначена для задач автоматической обработки текстов. Одним из приложений АвиаОнтологии является система автоматического анализа специальных технических текстов типа «Логика работы...» [2]. Задачи приложения определили специализацию АвиаОнтологии, которая связана, в первую очередь, с расширенным описанием области процессов обмена и передачи информации при решении задач авиационной тематики.

1 Структура АвиаОнтологии

АвиаОнтология представляет иерархическую сеть понятий. Для каждого понятия описано множество синонимичных текстовых входов – синонимический ряд понятия. Понятия организованы в сеть на основе определенной системы отношений, которая включает таксономические отношения ВЫШЕ-НИЖЕ, отношение ЧАСТЬ-ЦЕЛОЕ, отношения ассоциации (концептуальной зависимости), а также специальные отношения временных последовательностей. Объем АвиаОнтологии на начало 2004 г. составляет 1600 понятий и 4700 текстовых входов.

Помимо специальных терминов и понятий из авиационной области АвиаОнтология содержит большой пласт понятий общезначимой лексики, необходимый для использования ресурса в приложениях. Так, АвиаОнтология содержит такие понятия как АЛЬТЕРНАТИВА, БЕЗОПАСНОСТЬ, ВЗАИМОДЕЙСТВИЕ, ВРЕМЯ, КОЛИЧЕСТВО, МАКСИМУМ, МИНИМУМ и др.

Верхние уровни АвиаОнтологии фактически определяют множество семантических классов, к которым «привязаны» по отношению ВЫШЕ понятия онтологии. Наиболее представительными классами являются семантические классы «свойство», «техническое устройство», «процесс». Включение семантических классов в онтологию позволяет устанавливать семантический класс понятия на основе механизма вывода по иерархии отношения ВЫШЕ.

2 Архитектура и инструментальные возможности системы «OntoEditor»

Разработка больших информационных ресурсов является весьма ответственной и трудоемкой задачей и требует наличия инструментальных средств поддержки разработки. Проектирование информационных ресурсов типа онтологий на основе специализированной программной системы «OntoEditor» осуществляется на основе визуальной технологии, позволяющей эксперту или инженеру по знаниям вводить понятия, синонимические ряды понятий, связи между понятиями и отображать введенную информацию в графическом режиме.

Инструментальная система визуального проектирования «OntoEditor» является специализированной СУБД. Система предназначена для ручного редактирования онтологий, хранящихся в реляционной базе данных в определенном формате, а также обслуживания запросов пользователей и внешних программ.

Инструментальная система позволяет:

- добавлять, изменять и удалять отдельные записи БД;
- автоматически корректировать данные при удалении конкретных записей (например, удалять отношения стертого концепта);
- поддерживать ведение нескольких онтологий, в том числе смешанных (например, с общими списками типов отношений, классов, синонимов и др.);
- импортировать онтологии различных форматов данные из внешних баз данных (механизмы импорта разрабатываются на конкретную базу данных);
- вести обработку онтологий в табличной и графической формах;
- поддерживать иерархическую структуру классов;
- выделять по заданному фильтру определенные подмножества редактируемой онтологии;
- вести автоматическую статистику по объектам онтологии;
- осуществлять поиск цепочек отношений концептов с заданными свойствами;
- выводить требуемую информацию на печать или в текстовый файл в заданных формах отчета в кодировках ANSI и ASCII;
- обрабатывать внешние запросы с использованием механизма обмена DDE (Dynamic Data Exchange).

Система поддерживает многооконный интерфейс и снабжена развитой системой подсказок, а также механизмами поиска конкретной записи.

3 Проектирование онтологий

Проектирование онтологии включает этапы введения понятий, синонимических рядов понятия, отношений, построение классов понятий. В [1] описана технология введения новых понятий и отношений, выделены основные принципы проектирования онтологий.

Основная процедура введения нового понятия состоит из следующих шагов:

- на основе анализа текстовой коллекции выделяется слово или словосочетание, обозначающее важное понятие предметной области, которому присваивается по возможности понятное и однозначное имя. Например, при вводе понятия "крыло самолета" в принципе можно назвать его *крыло* – в авиационной области это слово однозначно, но лучше дать более полное название, которое не потеряет однозначность при расширении области или при включении данной онтологии в более обширную онтологию;
- при вводе понятия требуется ввести хотя бы одно отношение этого понятия с другими понятиями онтологии. Как правило, достаточно сложно сразу описать все необходимые отношения нового понятия. Ввод нового понятия часто влечет внесение уточнений в уже существующие описания, исправление неточностей описаний;
- понятие должно быть снабжено списком слов и словосочетаний, с помощью которых можно сослаться в тексте на вводимое понятие. В качестве таких текстовых входов могут быть отдельные слова (существительные, прилагательные, глаголы), а также именные и глагольные группы. При этом многословное языковое выражение должно употребляться в текстах как неразрывная конструкция. Если текстовый вход является многозначным, то он должен быть помечен как многозначный.

Инструментальная система «OntoEditor» поддерживает процесс введения новых понятий и отношений с помощью набора команд меню «Таблица».

Команда «Классы» позволяет выстраивать иерархию классов понятий.

Команда «Отношения» предназначена для редактирования типов отношений онтологии.

При определении типа отношения необходимо задать следующие атрибуты:

- *имя типа* отношения;
- *направление* (характеристика, определяющая направленность отношения между двумя концептами). В табличной форме направленность отношения выражается названиями концептов, связанных данным отношением: концепт-источник, концепт-приемник. Для графического представления отношения используется соответствующий вид связи. Атрибут «Направление» может принимать четыре значения: неопределенное, однонаправленное, двунаправленное, ненаправленное. В соответствии со значением атрибута «Направление» вводятся имена прямого и обратного отношения, основная цель которых установить однозначное соответствие между внутренним типом отношения и его внешним эквивалентом в импортируемой онтологии;
- *описание* – краткое описание сущности данного типа отношения.

Команда «Концепты» осуществляет редактирование концептов онтологии, выбранных с помощью механизма фильтрации. Для фильтрации используются следующие атрибуты концепта: *название, класс, синонимы и их количество, количество связей, абстрактность*.

Команда «Синонимы» предназначена для редактирования синонимов онтологии, выбранных с помощью механизма фильтрации. Для фильтрации используется атрибут *концепт*, по которому выбираются все синонимы заданного концепта.

Команда «Связи» позволяет редактировать отношений между концептами онтологии, выбранных с помощью механизма фильтрации. Для фильтрации используются следующие атрибуты: *название концепта-источника, название концепта-приемника, имя типа отношения*. Механизм фильтрации допускает просмотр логической суммы отношений, удовлетворяющих либо одному, либо другому атрибуту (флажок "ИЛИ"). Инструментальная система «OntoEditor» поддерживает развитый механизм поиска в онтологии по запросу пользователя. В системе реализован поиск концептов и отношений на основе механизма фильтрации, а также поиск цепочек связей заданной длины между двумя определенными концептами. Механизм поиска цепочек связей позволяет находить в онтологии цепочки заданной длины, а также цепочки в диапазоне указанных длин.

На рисунке 1 показаны в сравнении результаты поиска цепочек связей между концептами «Авария» и «Неисправность» при различных максимальных длинах цепочек (в окне слева максимальная длина цепочки связей равна 2, в окне справа – 5).

Механизм поиска цепочек связей между концептами позволяет исследовать структуру онтологии, предоставляя эксперту удобный интерфейс для тестирования структурных характеристик ресурса. При встраивании онтологии в прикладные системы, например в систему анализа технических текстов, основной задачей которой является контроль информационной целостности и полноты специальных тестов документации, механизм поиска цепочек связей является инструментом, позволяющим устанавливать наличие связей между концептами, а также исследовать узлы и тип построенной цепочки.

Типы запросов, обслуживаемых системой, постоянно расширяются исходя из потребностей исследования структуры онтологии. Команда «Печать» позволяет вывести на внешние носители результаты поиска в онтологии.

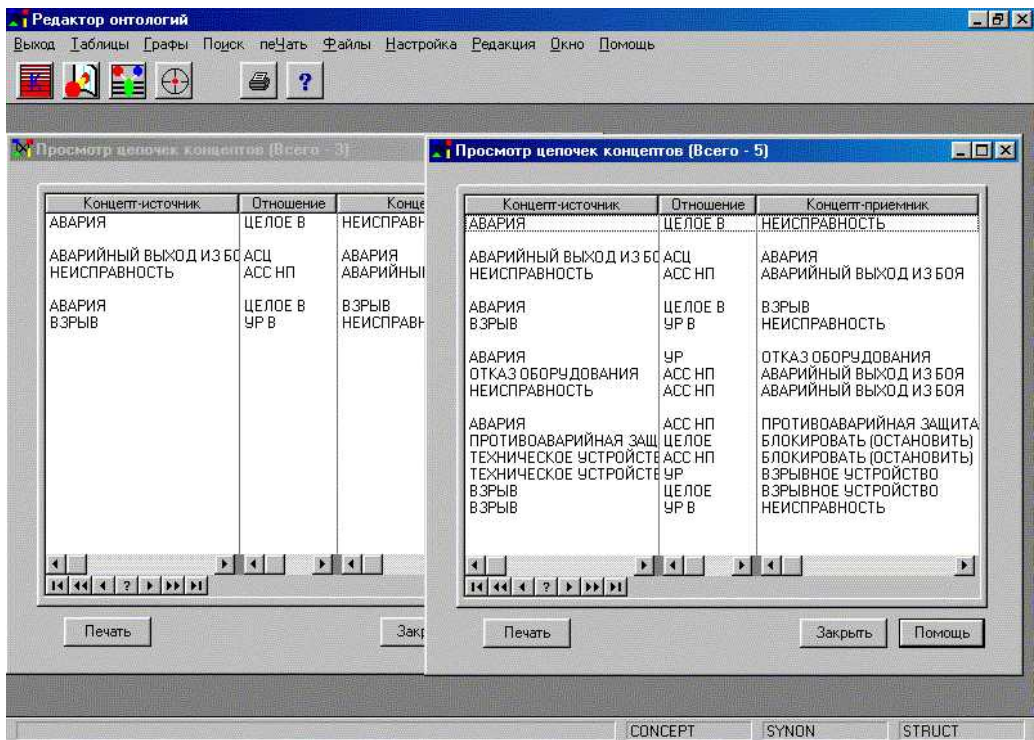


Рисунок 1– Поиск цепочек отношений между концептами.

4 Визуализация онтологии

Структура онтологии отображается на плоскости в виде графических образов концептов, отношений концептов и синонимов. Для визуализации онтологии используются следующие графические образы: концепты изображаются квадратами с названиями внутри, синонимы – окружностями (равномерно расположенными вокруг концептов), а связи – линиями. Тип отношений маркируется цветом линии, класс – цветом квадрата и текста.

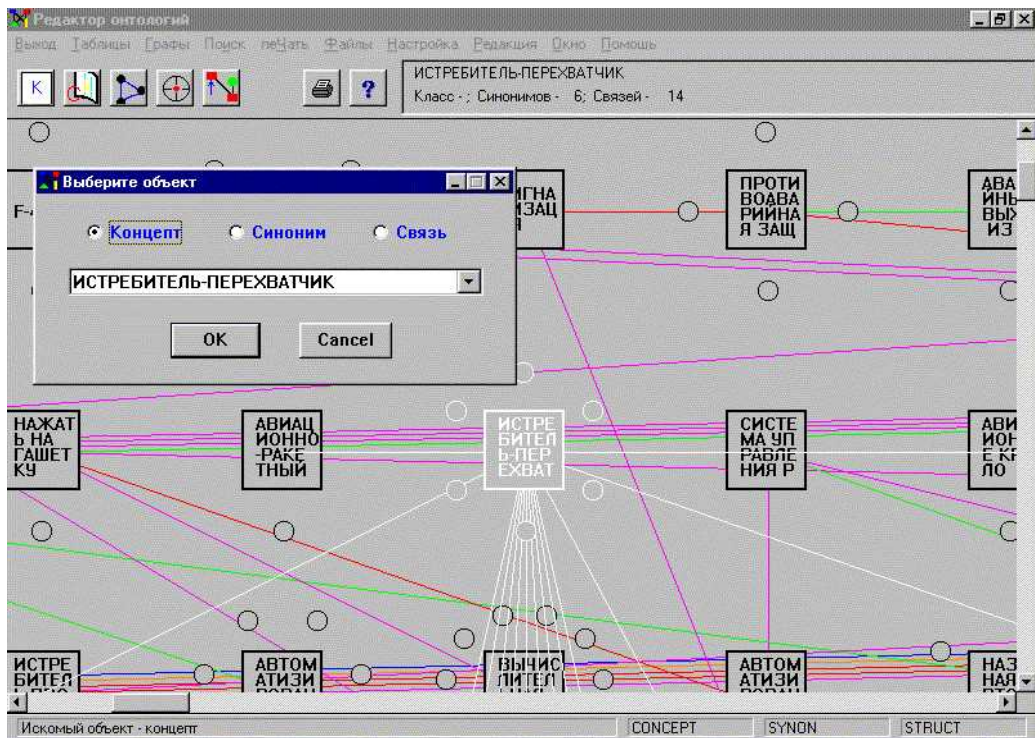


Рисунок 2 – Визуальный образ АвиаОнтологии

Визуальная работа над онтологией поддерживается следующими группами операций над одним или несколькими объектами одновременно:

- *масштабирования* (n-кратное увеличение или уменьшение изображения, переход к предыдущему виду, выбор полного обзора онтологии с минимально допустимыми размерами графических объектов, отображения на полный экран выделенной указателем мыши области или по границе зоны расположения селектированных образов);
- *центрирования* (графический результат операции располагается в центре экрана при выбранном режиме масштабирования, либо в центре экрана располагается участок изображения с координатами указателя мыши на момент выполнения операции);
- *ранжирования* (по порядку занесения в БД, по именам концептов, по классам, по количеству синонимов, по количеству связей, произвольно; выбор направления построения; автоматическое или ручное размещение при изменении онтологии);
- *фильтрации* (маскирования-демаскирования групп образов по различным критериям, например, выделения, отношения к тому или иному классу и др.);
- *выделения* (селектирования конкретных объектов или группы объектов для дальнейшей обработки);
- *поиска* (конкретных концептов, синонимов, отношений или цепочек отношений, существующих между двумя концептами) с последующим выделением;
- *драгирования* (перетаскивания графических объектов на плоскости при ручном режиме размещения);
- *информационного отображения параметров объектов* при прохождении указателя мыши над принадлежащей им "функциональной зоной".

Пользователю предоставляется возможность выбора:

- образов представления тех или иных объектов;
- цветовой гаммы представления;
- минимально возможных размеров графических образов;
- функций, закрепляемых за клавишами мыши и комбинациями их нажатия;
- выбора определенных образов для отображения (например, только концептов и связей без синонимов и названия концептов).

Разработан режим замораживания полученного видового экрана для работы с другими видовыми комбинациями обрабатываемой онтологии в многооконном режиме.

В настоящее время ведутся работы по поиску эффективных (с точки зрения полного анализа свойств онтологии) критериев упорядочивания и фильтрации объектов онтологии.

На рисунке 2 представлена визуальная форма АвиаОнтологии, на которой селектирован концепт ИСТРЕБИТЕЛЬ-ПЕРЕХВАТЧИК и множество его связей. Селектированный концепт имеет 14 связей (рассматривается степень данной вершины), которые заданы различным цветовым выделением. Окно выбора объекта, размещенное на форме позволяет осуществить переход к любому концепту АвиаОнтологии по имени.

Заключение

Визуальные методы проектирования онтологий способствуют более быстрому и более полному пониманию структуры знаний предметной области, что особенно ценно для исследователей, работающих в новой предметной области.

Инструментальная система «OntoEditor» предоставляет эффективный набор инструментальных средств, позволяющих осуществлять проектирование онтологии любой структуры, с любыми типами отношений и любыми классами концептов. Особенно важно, что система поддерживает импорт любого типа онтологии, реализованной на физическом уровне в табличной форме. Поддержка разнообразных поисковых запросов и механизмы вывода на онтологии позволяют исследовать внутреннюю структуру знаний предметной области. Система «OntoEditor» может быть встроена в приложения различного назначения, работающие с большими базами знаний.

Благодарности

Данное исследование выполнено при поддержке Российского Фонда Фундаментальных исследований, грант № 02-07-90279.

Список литературы

- [1] Dobrov B.V., Loukachevitch N.V., Nevzorova O.A. The Technology of New Domains' Ontologies Development // Proceedings of X-th Int. Conf. KDS-2003: "Knowledge-Dialogue-Solution", June 16-26, 2003, Varna, Bulgaria. P. 283-290.
- [2] Невзорова О.А., Федун Б.Е. Система анализа технических текстов «ЛОТА»: основные концепции и проектные решения.// Изв. РАН. Теория и системы управления. 2001. № 3. С. 138-149.

МОДЕЛИРОВАНИЕ РЕСУРСНЫХ ОГРАНИЧЕНИЙ ПРИ СТРУКТУРНОМ НЕДООПРЕДЕЛЕННОМ ПЛАНИРОВАНИИ В ТЕХНОЛОГИИ TIME-EX

И.Д. Гофман¹, Д.А. Инишев¹, А.В. Шурбаков¹, Л.Г. Романов²

¹Российский Научно-исследовательский институт искусственного интеллекта
Москва, 125190, а/я 85

hofman@aha.ru, dinishev@mtu-net.ru, andrshu@yandex.ru,

тел: (095) 155-45-30, факс: (095) 155-45-30

²Московский Государственный технический Университет гражданской авиации

Romanov@mstuca.ru

тел: (095) 459-64-52

Ключевые слова: интеллектуальная, недоопределенная, модель, планирование, ресурсы, структура, функциональная сеть, управление

Abstract

In the present paper, we examine some issues related to the resource constraints modelling in the structured subdefinite scheduling and planning (Time-EX system). The modelling based on the SD-models apparatus (Narin'yani, RRIAI). The more efficiency of the scheduling and planning has been postulated especially under conditions of high dimensionality of the plans and in the presence of resource constraints.

Введение

В последнее время исследовательские группы, занимающиеся развитием методов недоопределенной математики и вопросами приложения N-моделей ([1, 2]) активно прорабатывают вопросы структурных динамических N-моделей. Для проекта Time-EX эти проблемы весьма актуальны, т.к. по мере практического внедрения системы возникает необходимость в формировании планов большой размерности и введения компонент ресурсного планирования.

Комплекс работ проекта описывается сетью – ориентированным графом с двумя отмеченными вершинами – полюсами. Каждая вершина v сети N соответствует определенному событию в процессе выполнения работ из данного комплекса. Сетевая модель общеупотребительна и строится по методике «работа-связь», т.е. узлы направленного графа изображают задания (работы, операции) проекта, а ребра – отношения (связи) между заданиями.

Однако при больших размерностях плана (проекта с большим количеством заданий, операций) сетевая модель становится громоздкой, в которой затруднен поиск противоречивых ограничений (особенно ресурсных). Это приводит к несовместности модели и требует ее структуризации. При этом полученная структурная модель должна быть эквивалентна исходной по ее логической структуре, т.е. последовательность работ проекта (отношения между заданиями) должна оставаться неизменной.

Можно представить модель в виде совокупности модулей, связанных *деревом* транзитивных отношений *вложенности*. Т.о. каждый модуль может включать в себя произвольное число других модулей (в том числе, ни одного), но при этом каждый модуль может быть непосредственно вложен *только в один модуль*. Корень этого дерева – сама общая модель. Именно по этому пути ведется разработка новой реализации универсального решателя UniCalc ([4, 5]).

При этом структурная модель плана позволяет реализовать процесс удовлетворения ограничений на N-модели плана, т.к. множество ограничений, заданных на множестве переменных модели, являются либо элементарными, либо сами являются моделями (подмоделями), состоящими из локальных переменных и ограничений на них [3].

Более того, структурные модели позволяют реализовать динамику в Н-моделях, реализуя трансляцию ограничений из одной подмодели в другую.

1 Удовлетворение ограничений в структурной модели

Рассмотрим задачу удовлетворения ограничений в структурной Н-модели и передачу ограничений из одной подмодели в другую [3].

Мы называем *обобщенной вычислительной моделью* над Н-расширенными алгебрами $\langle *A, *F, *V, *C \rangle$ пару $M = \langle X, R \rangle$, где

$X = \{v_i \in V_{S_i} \mid i = 1, \dots, k\}$ – множество переменных,

$R = \{c_j \in C \mid j = 1, \dots, t\}$ – множество элементарных ограничений, связывающих переменные из X .

Вместо обычных универсумов X_v рассматриваются некоторые их недоопределенные расширения $*X_v$. Именно этот класс обобщенных вычислительных моделей называется недоопределенными моделями или Н-моделями.

Процесс вычислений на ОВМ имеет потоковый характер, - изменение объектных вершин сети активирует (вызывает к исполнению) функциональные вершины, для которых эти объектные вершины являются входными аргументами, а исполнение функциональных вершин в свою очередь может вызывать изменение результирующих объектных вершин. Вычисления заканчиваются тогда, когда либо не останется активных функциональных вершин (*УСПЕХ*), либо функция проверки корректности вырабатывает значение *ложь* (*НЕУДАЧА*).

В работе [3] доказаны следующие утверждения:

- 1) Процесс удовлетворения ограничений в Н-моделях завершается за конечное число шагов.
- 2) Достижение процессом *НЕУДАЧИ* или *УСПЕХА* предопределено входными данными (начальными Н-значениями переменных и ограничениями) и не зависит от конкретной стратегии выбора очередного ограничения для интерпретации.
- 3) В случае *УСПЕХА* процесса, при одних и тех же входных Н-значениях переменных их выходные Н-значения не зависят от конкретной стратегии выбора очередного ограничения для интерпретации.
- 4) В случае *УСПЕХА* процесса, решение задачи (если оно существует) лежит внутри полученного результата (декартова произведения Н-значений).

Согласно строгому определению Н-модели [3] каждое элементарное ограничение является монотонным отображением $\varphi_j : X \rightarrow X$. Процесс удовлетворения ограничений в случае *УСПЕХА* есть результат применения отображения $\varphi = \varphi_1 \circ \varphi_2 \circ \dots \circ \varphi_m$, т.е. отображение φ есть монотонное отображение как суперпозиция монотонных отображений. Из доказательства данной теоремы, задача удовлетворения конъюнкции ограничений может быть представлена как суперпозиция удовлетворения элементарных ограничений*. Это означает, что мы можем структурировать Н-модель, которая образуется из множества переменных и ограничений на них, каждое из которых может быть либо элементарным, либо само являться моделью (подмоделью) с элементарными ограничениями над ее локальными переменными.

Множество n параметров-переменных модели задает n -мерное пространство, являющееся декартовым произведением областей значений всех переменных данного множества. Точки этого пространства соответствуют n -мерным наборам значений всех переменных. Каждая модель определяет в пространстве значений своих параметров некоторое тело (назовем его *телом решений*), образованное точками, удовлетворяющими всем отношениям данной модели.

* При этом, по замечанию А.С. Нариньяни, элементарное ограничение может быть вложенным дизъюнктивным Н-отношением, т.е. быть, например, моделью вида: $\{X, \{\{ Y \dots \} \dots \{ \dots \} Z \} \dots \}$

Если модель полностью определена, ее тело решений состоит из единственной точки. Если она противоречива, тело решений – пусто. И, наконец, в общем случае, при недоопределенной системе отношений оно содержит множество точек или целые гипертела. Проекция тела решений на каждый из параметров задают их области значений для данной модели.

Добавление к модели M любого дополнительного ограничения определяет модель M' , тело решений которой, как очевидно, вложено в тело решений исходной модели M . Соответственно, области значений части или всех параметров внутри M' в общем случае уже, чем внутри M .

Если ограничение задается сужением области значений какого-либо из параметров, то из тела решений убираются части, проекция которых попадает в исключенную зону значений данного параметра. Например, при ограничении, сужающим интервал значений числовой переменной изменением одной из его границ, от тела решений исключаемая часть отсекается $(n - 1)$ -мерной гиперплоскостью, проходящей через новую границу интервала. Подобные ограничения являются типичными для многих приложений, представляя технические требования для трансформатора или финансовые условия для инвестиционного проекта.

Реализация вычислений на структурной Н-модели возможна несколькими путями. Один из них предложен в [4, 5]. В разрабатываемую версию универсального решателя введен *сетевой оператор*, который позволяет не только задавать систему ограничений, используя текстовое представление модели, но и строить функциональную сеть вручную, явным указанием узлов, их типов и значений, а также связей между узлами. При этом допускается возможность слияния нескольких функциональных сетей, хранящихся в разных экземплярах сетевых операторов, в одну сеть, а также свободное изменение уже построенной сети. Все это позволяет обойтись во многих случаях без компилятора.

Как видно, операции с функциональной сетью (ее построение и изменение) и собственно вычисления логически разделены. Сетью заведует сетевой оператор, а вычислениями – решатель. Это упрощает использование системы, а также ее разработку и развитие.

Проиллюстрировать этот процесс можно на примере задач планирования и управления проектами.

План некоторого проекта есть сеть заданий (работ), связанных между собой некоторыми отношениями [6, 7]. Каждое задание можно рассматривать как интервальное событие или t -интервал, ограниченный t -точками. Недоопределенный календарный план можно представить в виде Н-модели $\hat{M} = (\hat{X}, \hat{R})$, где \hat{X} – множество недоопределенных t -интервалов, представляющих сроки выполнения планируемых заданий, а \hat{R} – множество отношений, связывающих t -интервалы отдельных заданий.

Переходя к структурной модели плана, имеет смысл рассматривать модель интегрированного плана как глобальную, а модели, входящих в него подпланов, как локальные модели или подмодели.

Локальная модель – это часть модели, выделенная пользователем как подмодель в силу каких-либо причин. В нашем случае модель недоопределенного t -интервала $X = \{x, y, d\}$ может быть выделена в отдельный блок (назовем его A). Отношение $R = \{y_i + r_i + 1 = x_k, r \geq 0\}$ (отношение следования с лагом $r \geq 0$) также является локальной моделью B . Другой блок – B представляет собой ресурсное ограничение (ограничение по стоимости) – $Cr = \{l_b, v_b, c_i\}$, где v_i – условный объем работы для каждого задания.

Все подмодели являются функционально интерпретируемыми и могут быть представлены в виде Н-сети, которая может быть названа подсетью сети некоторого подплана, содержащего эти подмодели.

При рекомбинациях Н-сети появляется возможность добавления ограничений в модель или организовать передачу ограничений из одной подмодели в другую без ее полной перекомпиляции. При учете ресурсных или финансовых ограничений аналогичным образом может

быть реализовано выравнивание ресурсов или корректировка календарного плана для оптимизации финансовых показателей.

Н-отношение $*R (*a_1, \dots, *a_n)$ интерпретируется Н-функциями

$$*a_i = *f_i (*a_1, \dots, *a_{i-1}, *a_{i+1}, \dots, *a_n), \quad i = 1, \dots, n.$$

Если значение числовой Н-переменной представлять интервалом $[x_-, x^+]$, где x_- – минимум, а x^+ – максимум значения x , то уравнение нашего примера будет интерпретироваться шестью функциями интервальной алгебры [10]:

$$z^- := x^- + y^-; \quad z_- := x_- + y_-;$$

$$x^+ := z^- - y_-; \quad x_- := z_- - y^-;$$

$$y^+ := z^- - x_-; \quad y_- := z_- - x^-;$$

2 Элементы реализации

Рассмотрим модель простого плана, представленного на таблице заданий (рисунок 1) и сетевой диаграмме (рисунок 2), представляющей отношения между заданиями.

Кроме того, в модели присутствуют некоторые ресурсные ограничения. На каждое задание назначаются некоторые ресурсы. Стоимость этих ресурсов образует стоимость задания. Стоимости заданий затем суммируются, и определяется стоимость (бюджет) проекта в целом. Сроки и длительность заданий и всего проекта зависят от наличия и достаточности ресурсов. Недостаточность ресурсов увеличивает сроки выполнения и снижает экономическую эффективность проектов, с другой стороны, ресурсные ограничения все же предполагают и некоторое разумное уменьшение себестоимости проекта, и снижение издержек. В качестве простой иллюстрации ресурсные ограничения можно продемонстрировать введением только стоимостных ограничений.

Исходные оценки стоимости заданий вводятся в Time-EX (в графу Таблицы заданий «Планируемая стоимость», рисунок 2). В процессе расчета расписания вычисляется условный стоимостной «Объем» (т.е. стоимость мероприятия на его длительность) и фиксируется в модели. В процессе работы над планом мы можем оценить влияние величины финансирования на длительность того или иного мероприятия. Зависимость очевидная: чем ограниченнее финансирование, тем большая продолжительность работы. Большее финансирование позволяет привлечь дополнительные ресурсы и сократить время выполнения работы. Time-EX позволяет дать количественные оценки, причем определить границы необходимых величин финансирования, когда модель становится несовместной и план нереализуем. Это иллюстрируется на рисунке 3. На рисунке 4 исходная планируемая стоимость мероприятия 3 восстановлена и подплан рассчитан и сжат с сохранением приемлемых резервов времени. В нижней части диаграммы – окно гистограммы изменения величины средней стоимости (затрат) во времени выполнения.

Time-EX for Windows - [Окно таблиц и диаграммы Гантта: C:\Time-EX2_2_b_3_1\Documents\136_00.tme]

Файл Правка Задания Вид Проект Опции Окно Помощь

ID	Состояние	Имя	Длительность	Начало	Окончание	Объём	Общая стоимость
Проект	<input type="radio"/>	п/п 136	30-42	1-8	37-43	0	1839
1	<input type="radio"/>	Подготовительные мероприятия	12-18	2-8	13-19	4812-15238	401
2	<input type="radio"/>	Строительство	12-18	2-8	13-19	6720-13440	560
3	<input type="radio"/>	Основные мероприятия	12	14-20	25-31	660	55
4	<input type="radio"/>	Завершающие мероприятия	12-18	26-32	37-43	9876-19752	823

Рисунок 1 – Таблица заданий для примерного подплана

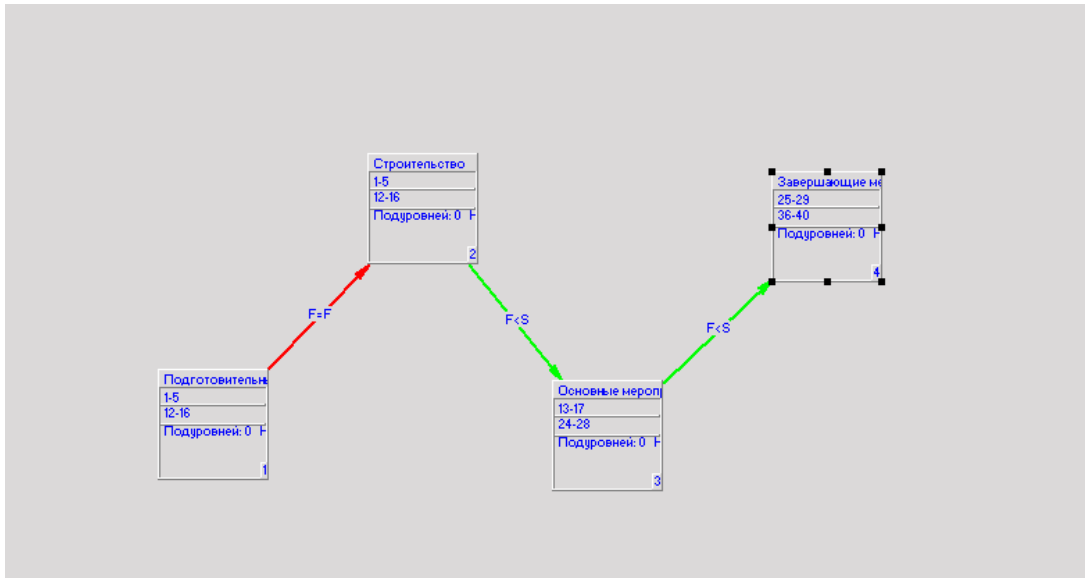


Рисунок 2 – Сетевая (PERT) диаграмма примерного подплана

Time-EX for Windows - [Окно таблиц и диаграммы Гантта: C:\Time-EX2_2_b_3_1\Documents\136_00.tme]

Файл Правка Задания Вид Проект Опции Окно Помощь

ID	Состояние	Имя	Длительность	Начало	Окончание	Объём	Общая стоимость
Проект	<input type="radio"/>	п/п 136	30-42	1-8	37-43	0	1738
1	<input type="radio"/>	Подготовительные мероприятия	12-18	2-8	13-19	4812-15238	300
2	<input type="radio"/>	Строительство	12-18	2-8	13-19	6720-13440	650
3	<input type="radio"/>	Основные мероприятия	12	14-20	25-31	660	20
4	<input type="radio"/>	Завершающие мероприятия	12-18	26-32	37-43	9876-19752	823

Ошибка
Inconsistent system
OK Справка

Рисунок 3 – Определение границ допустимого финансирования мероприятий

Заключение

В данной работе представлены некоторые вопросы, прорабатываемые в настоящее время в рамках проекта интеллектуальной технологии планирования Time-EX.

Модель плана с учетом проектной иерархии образует естественную структуру, которую удобно описывать, используя понятие функциональной сети. Вычислительное ядро системы осуществляет параллельный вычислительный процесс удовлетворения ограничений с управлением по данным, естественно распараллеливая вычисления. Данные обстоятельства могут позволить легко реализовать не только календарное, но и полномасштабное структурное ресурсное планирование в системе управления проектами, а также в системах производственного планирования.

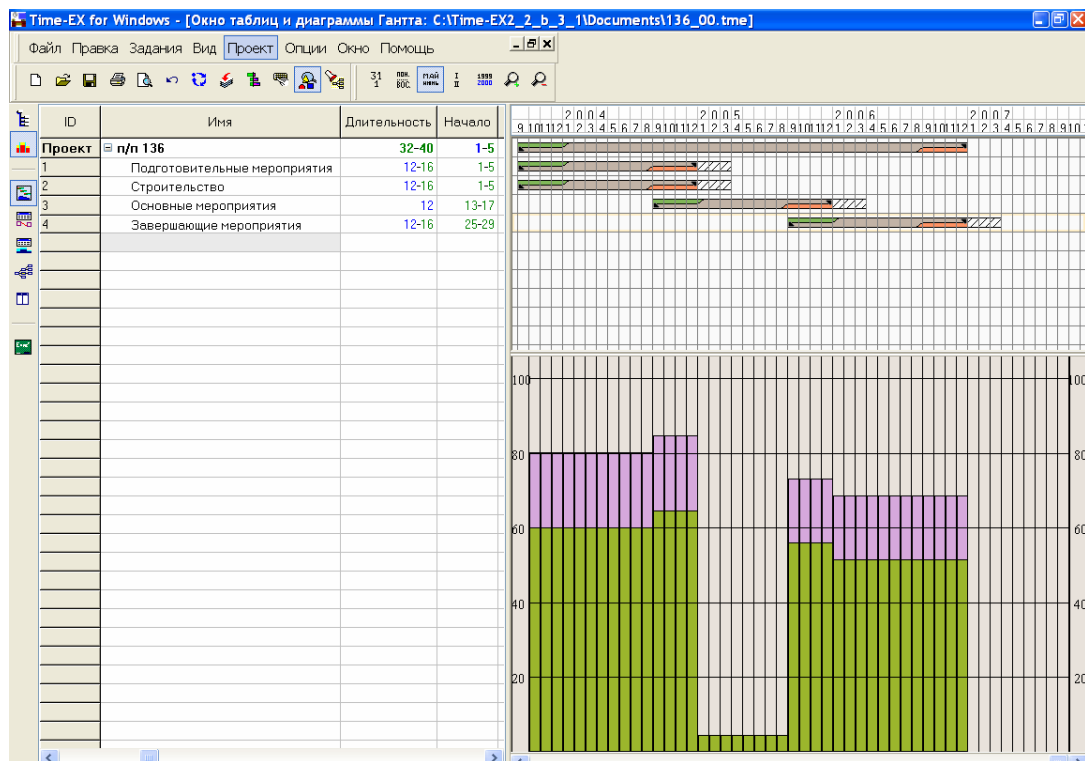


Рисунок 4 – Корректировка значений стоимости заданий и оптимизация подплана

Список литературы

- [1] Нариньяни А.С. Недоопределенность в системах представления и обработки знаний // Изв. АН СССР. Техн. кибернетика. – 1986, N5. – С. 3-28.
- [2] Нариньяни А.С., Семенов А.Л., Телерман В.В., Швецов И.Е., Яхно Т.М. Недоопределенные модели и их приложения // Системная информатика. Вып. 5: Архитектурные, формальные и программные модели. - Новосибирск: Наука. - 1996. - С.132-198.
- [3] В.В. Телерман, Д.М. Ушаков, Недоопределенные модели: формализация подхода и перспективы развития. Сб. Проблемы представления и обработки не полностью определенных знаний, Российский Научно-исследовательский институт искусственного интеллекта, Москва-Новосибирск, 1996, с.7-30.

- [4] Ю.А. Загорулько, В.О. Кошечев, П.Г. Мамонтов, Д.Ю.Парамзин, Новый подход к разработке решателя UniCalc, Труды 8-й национальной конференции по искусственному интеллекту - КИИ'2002, Москва, Физматлит, 2002, Т.2, с.720-728.
- [5] Ю.А. Загорулько, В.О. Кошечев, П.Г. Мамонтов, Д.Ю. Парамзин, Универсальный решатель задач с архитектурой, поддерживающей динамические недоопределенные вычисления, Сборник «Проблемы управления и моделирования в сложных системах». Труды 5-й международной конференции 17-20 июня 2003 года. Самара, Россия, 2003
- [6] Нариньяни А.С., Седреева Г.О., Седреев С.В., Фролов С.А., Time-EX|Windows – Новое поколение технологии календарного планирования, Сб. Проблемы представления и обработки не полностью определенных знаний, Российский Научно-исследовательский институт искусственного интеллекта, Москва-Новосибирск, 1996. – с.101-116.
- [7] Нариньяни А.С., Гофман И.Д., Инишев Д.А., Банасюкевич Д.В. Развитие интеллектуальной технологии недоопределенного планирования и управления проектами Time-EX, Труды II-ой международной конференции CSCMP-2000, Самара, 2000. с. 234-244.

НОВАЯ ВЕРСИЯ УНИВЕРСАЛЬНОГО РЕШАТЕЛЯ UNICALC: ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ РАЗВИТИЯ¹

Ю.В. Костов¹, Д.А. Липовой², П.Г. Мамонтов³, Е.С. Петров²

¹ЗАО “ИнтелиТек”
125190, Москва, а/я 85
kostov@iis.nsk.su
тел: +7 (095) 1550-45-30

²Российский НИИ Искусственного Интеллекта
630090, Новосибирск, пр. Лаврентьева, 6
multik@gorodok.net, pes@iis.nsk.su
тел: +7 (3832) 32-83-59, факс: +7 (3832) 32-83-59

³Институт Систем Информатики СО РАН
630090, Новосибирск, пр. Лаврентьева, 6
mamontov_p@ngs.ru
тел: +7 (3832) 34-29-91

Ключевые слова: недоопределённые вычисления, программирование в ограничениях, вычислительные модели

Abstract

UniCalc is a new generation software, based on the apparatus of subdefinite models, capable of solving arbitrary algebraic and logic constraints. The new version of the system has a number of changes that enhance the efficiency, as well as the quality of the computed solutions. In this paper we describe these new capabilities. Our attention is mainly focused on the symbolic transformations of constraints and on the correctness of the results of computations. We conclude by discussing the directions for future development of the system.

Введение

Традиционные методы решения систем алгебраических уравнений обычно основываются на специализированных алгоритмах вычислительной математики и применимы к задачам определенного класса (например, к системам линейных уравнений, полиномам и т.д.). Используемое их программное обеспечение позволяет решать только точно определенные задачи некоторого класса: под каждый класс задач – свой метод.

Однако, задачи, возникающие в реальной жизни, как правило, являются недо- или переопределенными, включают неравенства, логические условия, неточные значения, используют одновременно параметры различных типов. В такой ситуации методы классической вычислительной математики применимы лишь после (зачастую нетривиальных) преобразований и упрощений исходной задачи.

Решатель UniCalc реализует наиболее общий метод распространения ограничений [1, 2] и позволяет специфицировать задачу в виде, максимально приближенном к исходной её постановке. Кроме того, некоторые (или все) параметры задачи могут быть заданы неточно; вообще, точное значение есть частный случай неточного.

К настоящему моменту было создано несколько опытных версий решателя UniCalc [4, 5], различающихся по своим функциональным возможностям, реализующие те или иные архитектурные, технологические, вычислительные решения. Однако все эти прототипы, эксперимен-

¹ Работа выполнялась при финансовой поддержке РФФИ, проект № 03-01-00750.

тальные и опытные версии являлись лишь этапами длительного периода НИОКР и не получили широкого применения.

В докладе рассматривается версия решателя UniCalc нового поколения, которая, во-первых, базируется на 15-летнем опыте предыдущих разработок и включает в себя лучшее из того, что было сделано ранее; во-вторых, содержит ряд концептуальных новшеств, рассматриваемых ниже; в-третьих, имеет запроецированные точки роста в направлении средств спецификации моделей, тем самым расширяя область применимости решателя.

1 Краткий обзор аналогов

С математической точки зрения решатель UniCalc представляет собой программное обеспечение ЭВМ для надежного (т.е. гарантирующего корректность результатов) решения задач математического программирования. Все существующее программное обеспечение такого назначения реализует (в той или иной форме) методы интервальной математики. Перечислим его наиболее известные образцы.

Системы CLP(BNR), ECLIPSE, PROLOG-IV [6, 7, 8] являются интерпретаторами языка Пролог со встроенными интервальной арифметикой и алгоритмами распространения ограничений. Эти системы используются для решения задач дискретного математического программирования.

Для языков программирования С, С++, Фортран разработаны библиотеки *filib++*, PROFIL/BIAS, INTLIB, GlobSol, ILOG Solver [9, 10, 11, 12, 13], предоставляющие готовые реализации интервальных методов. Такие библиотеки используются при разработке специализированного математического обеспечения ЭВМ.

Интервальные методы встроены или реализованы для систем компьютерной алгебры: Matlab/IntLab, Maple/MuPad, Mathematica [14, 15, 16]. Такой подход ориентирован на использование методов интервальной математики в научных расчетах.

Интервалы целых и вещественных чисел и интервальная арифметика поддерживаются диалектами языков Pascal-XSC, С++ и Fortran 95 фирмы Sun Microsystems [17, 18, 19].

Системы Numerica, RealPaver [20, 21], UniCalc являются интерпретаторами языков, ориентированных на спецификацию задач математического программирования. Такие системы используются для выполнения надежных инженерных расчетов.

На стандартных тестовых задачах [22] UniCalc успешно конкурирует по эффективности с системами Numerica и RealPaver, ближайшими ей по назначению. Вместе с тем она предоставляет более богатый язык для спецификации задач и обеспечивает возможность решения задач с целочисленными переменными.

2 Решатель UniCalc

2.1 Концепция Н-вычислений

В данном разделе мы излагаем только основные идеи недоопределённых вычислений (Н-вычислений); подробнее о Н-вычислениях см. статьи [1-3].

Недоопределённым значением (Н-значением) переменной a называется множество её возможных точных значений. Искомые и вспомогательные параметры задачи, значения которых полностью или частично не определены (недоопределены), могут быть *адекватно* представлены Н-значениями соответствующих типов.

Отношение $R(x_1, \dots, x_n)$ называется *функционально интерпретируемым*, если найдётся множество f_1, \dots, f_n функций интерпретации, позволяющих вычислить каждый аргумент x_i отношения через его остальные аргументы: $x_i = f_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$.

Ограничением называется функционально интерпретируемое отношение. В решателе UniCalc аргументами ограничений являются Н-значения.

H-моделью называется конечное множество ограничений. H-вычисления применяются для решения задач, специфицированных в виде некоторой H-модели. Как показывает опыт, этот класс достаточно широк: ресурсное и календарное планирование, составление расписаний, САПР и пр.

Решением H-модели $M = \{R_1(x_1, \dots, x_n), \dots, R_m(x_1, \dots, x_n)\}$ называется набор значений (a_1, \dots, a_n) , удовлетворяющий всем ограничениям M : $R_1(a_1, \dots, a_n), \dots, R_m(a_1, \dots, a_n)$.

Во время решения задачи каждое H-значение представляет собой множество возможных значений некоторого параметра x_i . Сначала H-значение совпадает с доменом типа данных параметра x_i или является его подмножеством. В процессе решения H-значение автоматически *монотонно стягивается* за счет удаления значений, заведомо не являющихся искомыми значениями x_i . Если хотя бы одно H-значение модели становится пустым, это означает, что модель несовместна, т.е. задача не имеет решения.

Процесс стягивания H-значений реализуется алгоритмом недоопределённых вычислений (АНВ), который вместе с библиотекой математических функций образует ядро решателя UniCalc. В общих чертах, происходит это следующим образом.

Используя набор стандартных математических и логических операций, пользователь специфицирует задачу в виде H-модели. Компилятор решателя раскладывает ограничения H-модели на *элементарные ограничения*, каждое из которых содержит единственную стандартную операцию. Для всех стандартных операций реализованы функции интерпретации, которые образуют математическую библиотеку решателя.

Функции интерпретации полученных элементарных ограничений выстраиваются в очередь и запускаются на счет. Если функция интерпретации стягивает H-значение какого-либо параметра x_i , то в очередь добавляются все функции интерпретации, в которые параметр x_i входит в качестве аргумента. Таким образом, в АНВ осуществляется передача управления по данным (data driven). Цикл заканчивается, когда очередь пуста или получено пустое H-значение. При этом гарантируется [3]:

- 1) результат вычислений не зависит от порядка вычислений функций интерпретации;
- 2) вычисления завершаются за конечное число шагов.

Помимо АНВ, в решателе UniCalc реализованы механизм обжатия интервальных H-значений и алгоритм поиска корней (точных решений).

Результатом обжатия интервального H-значения является такое его подмножество, у которого левая граница совпадает с самым левым корнем, а правая – с самым правым. Обжатие позволяет отбросить некоторые области возможных значений, в которых нет решений, и которые не может автоматически исключить АНВ.

Алгоритм поиска корней реализует метод деления пополам (бисекция): H-значение разбивается на две равномошные части, после чего происходит запуск АНВ на каждой половине. Если на одной из половин система оказывается несовместной – эта половина отбрасывается, иначе выполняется следующий шаг бисекции. Таким образом, рекурсивно, происходит спуск к каждому корню.

2.2 Архитектура и принцип работы

Новая версия решателя UniCalc имеет модульную архитектуру, которая представлена ниже на Рисунке 1. В решателе выделяется четыре основных программных блока: сетевой оператор, вычислитель, компилятор и менеджер, содержащий интерфейсные модули для взаимодействия с внешними системами.

Основной структурой данных в решателе UniCalc является *функциональная сеть*, в которой хранится модель в виде ориентированного ациклического графа. Переменные и константы изображаются листьями этого графа, элементарные ограничения – внутренними вершинами. Такое представление модели удобно для символьных преобразований.

Компилятор преобразует текстовое описание модели на языке UniCalc в функциональную сеть.

Сетевой оператор – это модуль, предоставляющий набор операций для работы с функциональной сетью. Он позволяет строить функциональные сети, как с помощью компилятора, так и без него – последовательно создавая вершины (операции, переменные, константы) и связывая их между собой.

Для уменьшения размеров функциональной сети и увеличения скорости работы всей системы в целом – применяется набор символьных преобразований сети (оптимизация). В текущей версии решателя реализованы следующие виды оптимизации: раскрытие скобок, вычисление константных выражений, замена известных переменных на константы, уменьшение количества линейных слагаемых в выражениях путем создания матрицы линейной части задачи и трансформации ее к треугольной форме и др.

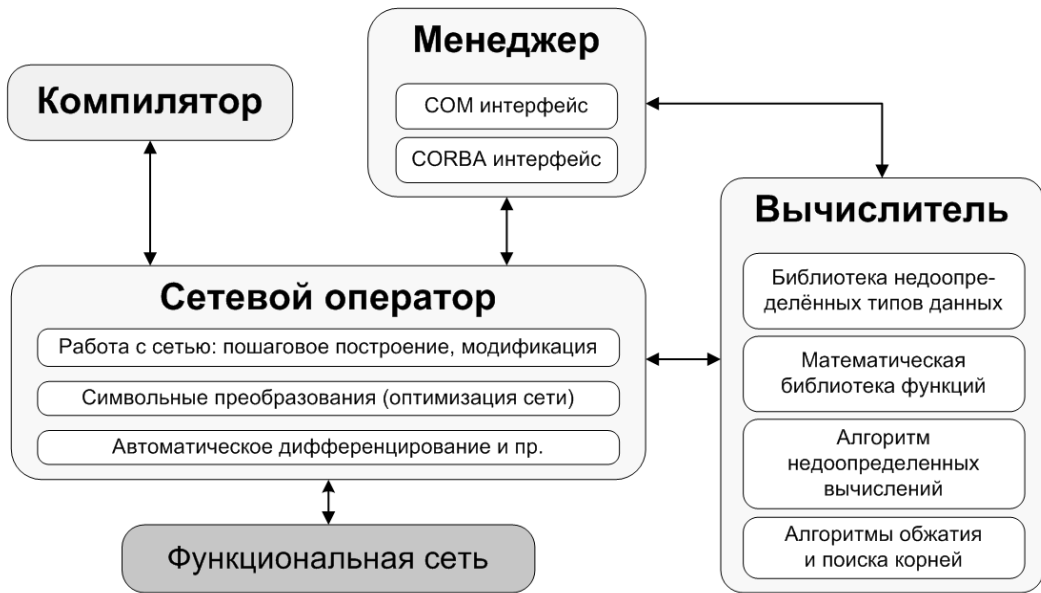


Рисунок 1. Общая архитектура решателя.

Вычислитель предназначен для выполнения следующих операций над системой ограничений, представленной функциональной сетью:

- 1) Вычисление значения выражения с недоопределёнными коэффициентами при помощи библиотеки интервальных функций интерпретации.
- 2) Использование АНВ для получения более точных оценок параметров модели.
- 3) Поиск корней, обжатие (минимизация ширины интервалов переменных, в которых лежат корни системы), максимизация/минимизация значений переменных.
- 4) Сервисные функции для разбиения глобальной вычислительной модели на более мелкие локальные вычислительные модели — блоки. Поддерживаемый вычислителем механизм манипулирования глобальной моделью позволяет динамически создавать и уничтожать блоки, устанавливать и изменять связи между ними. Кроме того, он обеспечивает средства для модификации содержимого блока, заключающиеся в изменении недоопределённых значений, связываемых данной вычислительной моделью.

Может быть создано произвольное число вычислителей и сетевых операторов.

В задачи менеджера входит управление сетевыми операторами и вычислителями.

2.3 Символьные преобразования

Символьные преобразования модели осуществляются на функциональной сети и имеют две цели: повышение точности вычислений и уменьшение времени счёта на больших моделях. Все преобразования условно разбиты на три группы:

- Встроенные – вычисление и удаление константных подвыражений, инициализация переменных и т.п.
- Базовые – приведение подобных слагаемых, раскрытие скобок и пр.
- Преобразования линейных ограничений (алгоритм Гаусса).

Рассмотрим некоторые виды преобразований более подробно.

Зачастую модель содержит константные выражения, значение которых не зависит от значений переменных (например, выражения без переменных, единица в некоторой степени и т.п.). Такие выражения заменяются соответствующими константами.

Эффективными символьными преобразованиями являются приведение подобных слагаемых и вынесение за скобки общего множителя. Такие преобразования уменьшают объем вычислений и одновременно повышают их точность.

В любой модели имеется одно или более ограничений сравнения: равенство, нестрогие неравенства. Функции интерпретации таких ограничений не производят иных действий, кроме пересечения аргументов этого ограничения с их новыми вычисленными значениями. Поскольку константные аргументы ограничений сравнения не меняются в процессе вычислений, то возможно одноразовое пересечение с константным аргументом на этапе символьных преобразований. После чего данное ограничение можно исключить из сети. Для моделей с большим числом уравнений/неравенств выгода может оказаться весьма ощутимой. Кроме того, вычисляя такие ограничения, мы можем обнаружить противоречие на ранних стадиях решения системы.

Отождествление подвыражений – ищутся одинаковые части функциональной сети, оставляется только один экземпляр, дубликаты выбрасываются, дуги перестраиваются на оставшийся экземпляр. Модель становится более компактной и лучше связанной.

Одним из слабых мест алгоритма N-вычислений являются линейные задачи, поскольку линейные зависимости влекут слабую сходимости АНВ. Для улучшения показателей счёта (время и точность) необходимо уменьшить количество линейных слагаемых. Это реализуется при помощи выделения линейной части системы и преобразования ее по алгоритму Гаусса – приведение матрицы к трапециевидной форме.

2.4 Обеспечение математической корректности результатов

Набор N-значений (A_1, \dots, A_n) включает набор значений (a_1, \dots, a_n) , если выполняются n соотношений: $a_1 \in A_1, \dots, a_n \in A_n$.

Набор N-значений является *корректным относительно модели M* (далее просто: *корректным*, когда модель фиксирована), если он включает каждое решение модели M . Например, решениями модели $M = \{x_1 + x_2 = 7, x_1 * x_2 = 6\}$ являются наборы вещественных чисел $(6, 1)$ и $(1, 6)$. Набор N-значений $([-10,10], [-10,10])$ является корректным, так как он изображает оба решения M . Набор $([-10,1], [-10,10])$ не является корректным, так как он не изображает решение $(6, 1)$.

Как известно, числа с плавающей точкой представляются в памяти ЭВМ в формате $m * 2^e$, где m и e – целые числа, лежащие в определённых диапазонах. Вещественные числа, которые не могут быть представлены в таком виде (например, число 0.6), в машинной арифметике представляются с ошибкой (это касается также и промежуточных результатов вычислений). Такие ошибки называются ошибками округления. Из-за отсутствия систематического учета ошибок округления предыдущие версии решателя UniCalc теряли решения некоторых моделей, т.е. выдавали некорректные N-значения или ошибочно сообщали об отсутствии решений.

Данная версия решателя является первой версией, учитывающей ошибки округления и *гарантирующей корректность* получаемых результатов.

Во время N -вычислений N -значения параметров модели изменяются только функциями интерпретации элементарных ограничений. Сохранение корректности набора N -значений параметров модели гарантируется надлежащей реализацией каждой отдельно взятой функции интерпретации каждого элементарного ограничения.

Реализация функций интерпретации элементарных ограничений определяется точностью и набором команд, которые обеспечивает математический сопроцессор (эмулирующая его библиотека). Решатель UniCalc рассчитан на математические сопроцессоры (эмулирующие его библиотеки), поддерживающие формат IEEE 754-1985 для чисел с плавающей точкой. Эти сопроцессоры позволяют управлять направлением округления результатов арифметических операций и вычислять значения функций \exp и \log с точностью 1ULP (единица в последнем двоичном разряде мантииссы), а значения функций \sin и \tan – с точностью 2ULP. Данным требованиям удовлетворяют, например, процессоры семейства Intel Pentium.

Заключение

Дальнейшее развитие решателя UniCalc в основном будет нацелено на повышение уровня средств спецификации задач, что, таким образом, расширит область применимости решателя. Кроме того, планируются работы по введению дополнительных средств для повышения удобства работы с решателем.

Более конкретно, развитие будет осуществляться в следующих направлениях.

- Введение нового вида недоопределённости: перечислимый. Перечислимое N -значение будет задаваться перечислением возможных точных значений, а не граничными условиями. Это, в свою очередь, позволит ввести перечислимый символьный тип данных.
- Введение округляемого вещественного типа данных – для финансовых расчётов.
- Добавление возможности описывать экстенциональные (табличные) ограничения.
- Диагностика причин возникновения противоречий.
- Введение в язык спецификации модели управляющих директив.
- Визуализация динамики вычислений – возможность проследить изменение выбранных параметров. Графическое отображение результатов вычислений.
- Введение дополнительных эвристик поиска точных решений. Динамическая смена эвристики в процессе вычислений.
- Полноценная поддержка ограничений дизъюнкции и импликации, а также введение элементов императивности для управления процессом вычислений.
- Поддержка распределённых вычислений.

Список литературы

- [1] Нариньяни А.С. Недоопределённость в системах представления и обработки знаний // Изв. АН СССР. Техн. кибернетика. – 1986, №5. – С. 3-28.
- [2] Нариньяни А.С., Семенов А.Л., Телерман В.В., Швецов И.Е., Яхно Т.М. Недоопределённые модели и их приложения // Системная информатика. Вып. 5: Архитектурные, формальные и программные модели. - Новосибирск: Наука. - 1996. - С.132-198.
- [3] Ushakov D. Some Formal Aspects of Subdefinite Models. - Novosibirsk, 1998. - 23 p. - (Prepr. / Siberian Division of Russian Acad. Sci. IIS; № 49).
- [4] Babichev A.B., Kadyrova O.B., Kashevarova T.P., Leshchenko A.S. and Semenov A.L. UniCalc, A Novel Approach to Solving Systems of Algebraic Equations. Interval Computations, No.2, 1993, pp. 29-47.
- [5] Загорюлько Ю.А., Кошечев В.О., Мамонтов П.Г., Парамзин Д.Ю. Универсальный решатель задач с архитектурой, поддерживающей динамические недоопределённые вычисления // Труды V-й международной конференции "Проблемы управления и моделирования в сложных системах" -Самара: Самарский Научный Центр РАН, 2003. -С. 261-265.

- [6] Benhamou F., Older W. Programming in CLP(BNR) // Proc. PPCP'94, Newport, USA, 1994.
- [7] Петров Е.С. Интервальная библиотека: опыт интеграции логического программирования и недоопределенных моделей // Программирование, №4, С.40-49, 1998.
- [8] Benhamou F, Bouvier F., Colmerauer A., Garreta H., Gilletta B., Massat J.-L., Narboni G. Le manuel de Prolog IV // — PrologIA: 1996.
- [9] Lerch M., Tischler G., Wolff v. Gudenberg J., Hofschuster W., Kramer W. filib++, a Fast interval library supporting containment computations. // Preprint 2003/4, – Universitat Wuppertal: 2003.
- [10] Knueppel O. PROFIL/BIAS — a fast interval library // J. Computing, -V.53, -N.3-4, -P. 277-287, 1994.
- [11] Kearfott R. B., Dawande M., Du K.-S., Hu C.-Y. Algorithm 737: INTLIB: a portable FORTRAN-77 interval standard function library // ACM Transactions on Mathematical Software, -V.20, -N.4, -P.447-459, 1994.
- [12] Hu C.-Y., Corliss G.F., Kearfott R.B., Schulte M., Stadtherr M.A. GlobSol project // <http://www.mscs.mu.edu/~globsol/>
- [13] <http://www.ilog.com/products/solver/>
- [14] Rump S.M. INTLAB — INTerval LABoratory // Csendes T. (editor) Developments in Reliable Computing, -P.77-105, -Kluwer: 1999.
- [15] Gerhard J., Oevel W. MuPAD tutorial // -Springer: 2000, -ISBN 3540675469.
- [16] Keiper J.B. Interval arithmetic in mathematica // The Mathematica Journal, -V.5, -P.66-71, 1995.
- [17] Клатте Р., Кулиш У., Негра М., Рац Д., Ульрих Х. Pascal XSC – язык численного программирования. // -ДМК Пресс Москва: 2000, ISBN 5940740251.
- [18] Sun Microsystems, Inc. C++ User's Guide // -iUniverse: 2004, ISBN 0595286526.
- [19] Sun Microsystems, Inc. Fortran User's Guide // -iUniverse: 2003, ISBN 059573202X.
- [20] Van Hentenryck P., Michel L., Deville Y. Numerica: a modelling language for global optimization // -The MIT Press:1997.
- [21] Granvilliers L. On the combination of interval constraint solvers // Reliable Computing, -V.7 -P.467-483, 2001.
- [22] Jansson C., Knuppel O. A global minimization method: the multi-dimensional case // Technical report 92.1, -TU Hamburg-Harburg: 1992. Новосибирск, 2000. – 58 с.

ALIX: МЕТОД ГРАФИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРОЦЕССОВ В СЛОЖНЫХ ДИСКРЕТНЫХ СИСТЕМАХ

А. Артиба¹, А. Донтэн¹, С.И. Ясиновский²

¹Facultés Universitaires Catholiques de Mons
Chaussée de Binche, 151 Mons BELGIUM
artiba@fucam.ac.be, dontaine@fucam.ac.be

²МГТУ им. Н.Э.Баумана
107005, Москва 2-я Бауманская, 5
iassinov@fucam.ac.be

Ключевые слова: моделирование процессов, реорганизация процессов, имитация, сложные дискретные системы, принятие решений

Abstract

This paper presents a graphical enterprises modeling method oriented to process simulation. Among static, decisional and dynamical contexts of modeling, only the dynamical context is treated in this article. This formalism was inspired by well-known methods like SADT and UML. It is a part of a larger analysis and optimization environment that is also in development in our research centre (the CREGI). The goal of this environment is to simplify the use of simulation and optimization methods in industrial contexts.

Введение

В последние десятилетия роль моделирования сложных дискретных систем (в частности предприятий) существенно возросла. Различные модели стали неотъемлемой частью процесса проектирования и управления такими системами [1]. Современная конкурентная среда требует все более полного владения информацией и знаниями, выраженными, в частности, в виде различных формальных моделей [2]. Это привело к появлению многочисленных методов описания систем и сделало задачу выбора подходящего метода неочевидной. Вернада [3] рассматривает трудности выбора и использования современных методов моделирования конечными пользователями, имея в виду временные и бюджетные ограничения.

Ориентированное на имитационное моделирование описание сложной дискретной системы (СДС) является непростой задачей вследствие высокой сложности описываемых процессов (потoki элементов системы, ресурсы, ограничения, принятие решений, ...). Одним из направлений деятельности исследовательского центра CREGI (Centre de Recherches et d'Etudes en gestion Industrielle) католического университета Монса FUCaM (Facultés Universitaires Catholiques de Mons) является разработка инструментальной платформы, облегчающей и интегрирующей все этапы анализа и синтеза СДС (рисунок 1) на основе применения одного и того же набора концепций на различных уровнях процесса описания и решения, от методологии представления до решателей [4, 5]. Данный проект нацелен, в частности, на разработку формальных моделей, доступных для понимания всех специалистов, вовлеченных в процесс анализа, и в то же время адаптированных для автоматического преобразования во входные языки средств оптимизации и имитации. Эти модели должны позволять описывать статические и динамические аспекты систем, а также аспекты принятия решений. В рамках этой платформы графические модели служат связкой между предметной областью и методами решения. Частью этих работ является разработка языка графического описания дискретных процессов, ориентированного на их имитацию и оптимизацию. Этот язык должен быть достаточно полным и выразительным, чтобы представить информацию, необходимую для последующего применения формальных алгорит-

мов оптимизации и имитации специалистом-исследователем. С другой стороны, метод должен быть прост и интуитивно понятен, чтобы модели на его основе могли быть созданы и/или проверены специалистом, имеющим знания о моделируемой системе и процессе.

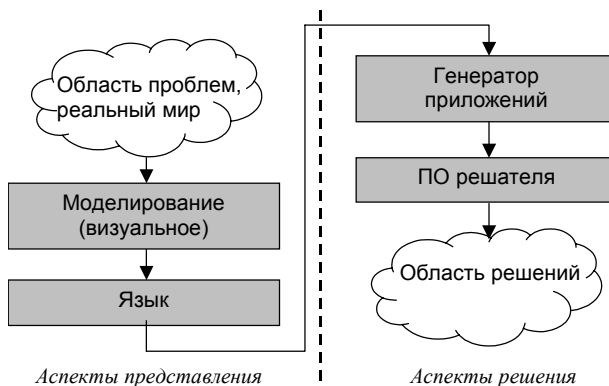


Рисунок 1 – Общий подход к решению проблем анализа СДС.

1 Анализ методов описания СДС

Существует большое количество методов и подходов к моделированию СДС. Назовем такие наиболее известные методы (в порядке их появления) как SADT, CIMOSA, GIM, ARIS, l'ENV40003, AIOSCOP, PERA et GERAM [6]. На основе этих методов были созданы многочисленные языки моделирования предприятий, в частности CIMOSA process language, ARIS, IDEF, IDEF*, IEM, WfMC, ODP, PSL/NIST,... Однако, несмотря на изобилие языков, трудно найти ПО моделирования, адаптированное для применения методов исследования операций. Немногочисленные коммерческие инструментальные средства предназначены либо для укрупненного описания СДС, либо для описания частных аспектов, либо не имеют встроенных средств имитации или методов оптимизации. С другой стороны, визуальные языки имитационного моделирования, такие, например, как Visual SLAM, часто имеют ограниченные выразительные возможности и концептуально весьма далеки от понятий предметной области, что весьма затрудняет овладение ими специалистами предприятий.

Проведя анализ ряда методов моделирования, мы пришли к выводу, что методами, наиболее отвечающими с концептуальной и практической точек зрения потребностям предприятий и других СДС, являются SADT и UML. Однако и эти методы не отвечают в полной мере требованиям, сформулированным выше.

Графические нотации SADT были созданы для моделирования функциональной и информационной структуры предприятий. Как отмечает Ким [7], в этом и сила, и слабость подхода. В частности, нотации SADT лучше подходят для описания (с различных точек зрения) того, что СДС (такая как предприятие) должна производить. Они не подходят для детального описания поведения различных элементов СДС (персонал, машины, программы и т.д.), необходимого для решения таких задач как, например, распределение ресурсов. Модели SADT/IDEF0 обеспечивают целостность моделируемой СДС на каждом уровне декомпозиции. Однако они статичны, они только представляют действия СДС и их взаимосвязи без прямого описания логических зависимостей и длительности [8]. Наконец, диаграммы SADT не описывают способ преобразования состояния ресурсов, выполняющих функцию, и элементов, над которыми выполняется функция [9].

UML является языком, но не является методологией моделирования. Он не предлагает четкой методики разработки моделей. Это привносит дополнительную гибкость, но в то же время затрудняет использование языка не экспертом, поскольку язык содержит мало семантической информации. Использование UML позволяет лучше владеть аспектами разработки СДС, предвидеть возможные варианты. Язык помогает обеспечить взаимодействие между различными командами, работающими над масштабным проектом. В [10] проводится анализ UML и утверждается, что этот язык не отвечает ряду ключевых принципов моделирования. Во-первых, не выполняется принцип единства, так как диаграммы последовательностей и кооперации семантически эквивалентны. Далее, есть проблемы с принципом консистенции, так как одна из диаграмм UML может противоречить другой. Невозможно составить диаграмму, содержащую вместе активности, объекты в них участвующие и их поддерживающие, что необходимо для имитации. Наконец, UML сложен в использовании.

В таблице 1 представлено обобщение результатов анализа SADT и UML (частично на основе [11, 12]) на основе критериев, выработанных в CREGI. Суммируя, можно сказать, что UML сложен и не всегда однозначен, а SADT прост и целостен, но не дает возможности описать ряд аспектов СДС, необходимых для формального преобразования созданных моделей в имитационные.

Таблица 1 – Сравнительный анализ методов моделирования.

	UML	SADT	Требования CREGI
Простота овладения	сложно	просто	просто
Графическое представление	да	да	да
Читаемость	слабая	хорошая	хорошая
Представление ресурсов	по требованию	упрощенное	детальное
Динамические аспекты (предусловия, состояние)	да	нет	да
Представление процесса	да	среднее	да
Динамика процессов	да	нет	да
Представление правил принятия решений	да	на уровне активности	да
Временные аспекты	да	упрощенно	да
Процессная ориентация	нет	да	да
Объектная ориентация	да	нет	да
Адаптированность к имитации	слабая	средняя	полная
Методология моделирования	нет	да	да
Область применения	широкая	ограниченная	широкая

2 Описание метода моделирования ALIX

Метод предназначен для моделирования процессов в СДС (в различных организациях и социо-экономических системах), прост и интуитивно понятен в использовании и ориентирован на создание имитационных моделей описываемых процессов. Основные элементы метода – активности, объекты и их потоки (связи).

Активности. Активность есть элементарный неделимый процесс, описываемый одним или двумя событиями. Ее выполнение мобилизует определенные ресурсы, может занимать некоторое время и имеет целью преобразование состояния участвующих объектов из входного в выходное. Активность характеризуется именем, входящими и выходящими объектами, объектами, ее выполняющими и иерархическим номером. Имеется три типа активностей и метаактивности (рисунок 2).

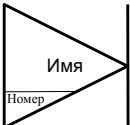


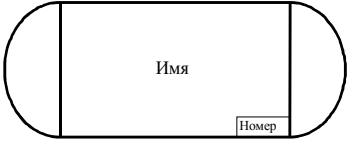
<p>Нерегулярное событие : Активность выполняется через определенные (заданные или случайные) интервалы времени независимо от состояния системы (одно событие).</p>	
<p>Правило : Активность, связанная с принятием решения; она выполняется при определенных условиях, изменяет состояние системы но не имеет длительности (одно событие).</p>	
<p>Действие : Активность, выполняющаяся при определенных условиях и преобразующая состояние системы в начале и в конце. Имеет длительность (два события).</p>	
<p>Метаактивность : Графический формализм, позволяющий представить в укрупненном виде некий подпроцесс и обеспечивающий таким образом иерархическую декомпозицию схемы процесса.</p>	

Рисунок 2 – Типы активностей в методе ALIX.

Объекты. Объект есть представление элемента СДС. Определено два типа объектов:

- «преобразуемые» объекты – те, которые подвержены описываемому процессу в СДС и которые меняют свое состояние по мере следования процессу;
- «используемые» объекты – те, которые используются для выполнения активностей.

Разделение это условно и зависит от контекста, оно введено для большей выразительности модели. Одни и те же объекты могут быть преобразуемыми и используемыми в различных процессах одной и той же СДС.

Объект представляется стрелкой со следующей информацией: имя объекта и имя класса объекта. Используемые объекты представляются вертикальными стрелками, направленными к активности и находящимися под ней. Для каждого объекта, вовлеченного в выполнение активности, указываются предусловия и могут быть указаны изменения состояния (рисунок 3).

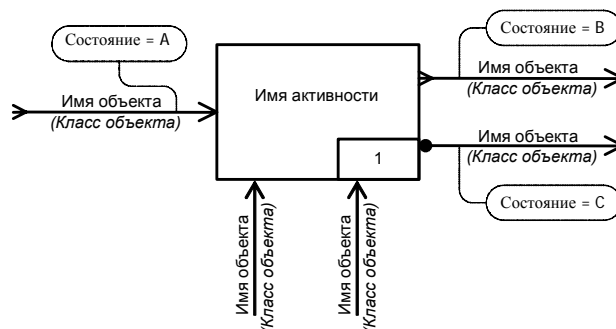


Рисунок 3 – Общее представление активности.

Связи. Связь между парой активностей представляет собой путь преобразуемого объекта между активностями. Объект покидает предшествующую активность в некотором состоянии, делая возможным выполнение следующей активности (рисунок 4). Связи также отображают информацию о создании и уничтожении объектов (рисунок 5).

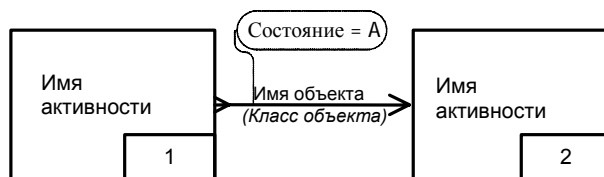


Рисунок 4 – Простая связь между двумя активностями.

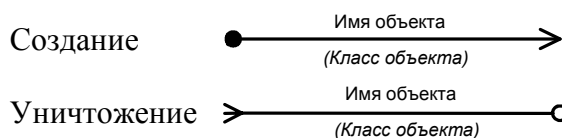


Рисунок 5 – Представление создания и уничтожения объектов.

Разновидностями простых связей являются ветвления и слияния. Метод также позволяет описывать циклы. Некий объект может подвергаться активности несколько раз, до тех пор пока его состояние не перестанет удовлетворять предусловию активности. Чтобы избежать заикливания, активность должна изменять состояние объекта таким образом, чтобы рано или поздно предусловие активности, связанное с данным объектом, перестало выполняться. Наконец, для облегчения схем с большим количеством связей, введены разрывы связей. Для нахождения продолжения прерванной связи используются символы разрыва и нумерация разрывов.

Метаактивность. Метаактивность есть способ группирования активностей для упрощения простоты и ясности модели. Таким образом, метаактивность описывает подпроцесс более низкого уровня, ее содержимое раскрывается отдельно. Метаактивности позволяют создавать иерархические модели (рисунок 6).

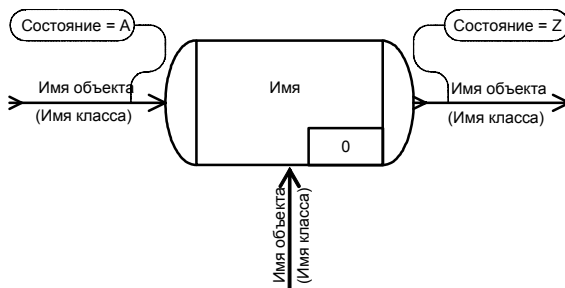


Рисунок 6 – Метаактивность в ALIX.

В настоящее время для тестирования и проверки подхода разработан прототип инструментального средства моделирования. Для быстрой разработки прототипа было использовано ПО VISIO фирмы Microsoft, которое просто в использовании, содержит все необходимое для манипуляции графическими объектами, позволяет экспортировать разработанные диаграммы в базу данных и разрабатывать программные модули на языке Visual Basic, что необходимо для генерации имитационной модели на основе визуальной.

3 Приложение

Разработанный метод моделирования прошел проверку в рамках нескольких реальных проектов по анализу и реорганизации производства. Опыт показал, что метод легко понимаем

и без каких-либо трудностей был принят специалистами предприятий как средство описания и обмена информацией.

На рисунке 7 представлен фрагмент модели реального процесса. Модель была разработана в течение нескольких дней. Она содержит десяток страниц, детально описывающих процесс производства и потоки материалов и документов в системе. На основе данной модели была разработана имитационная модель предприятия.

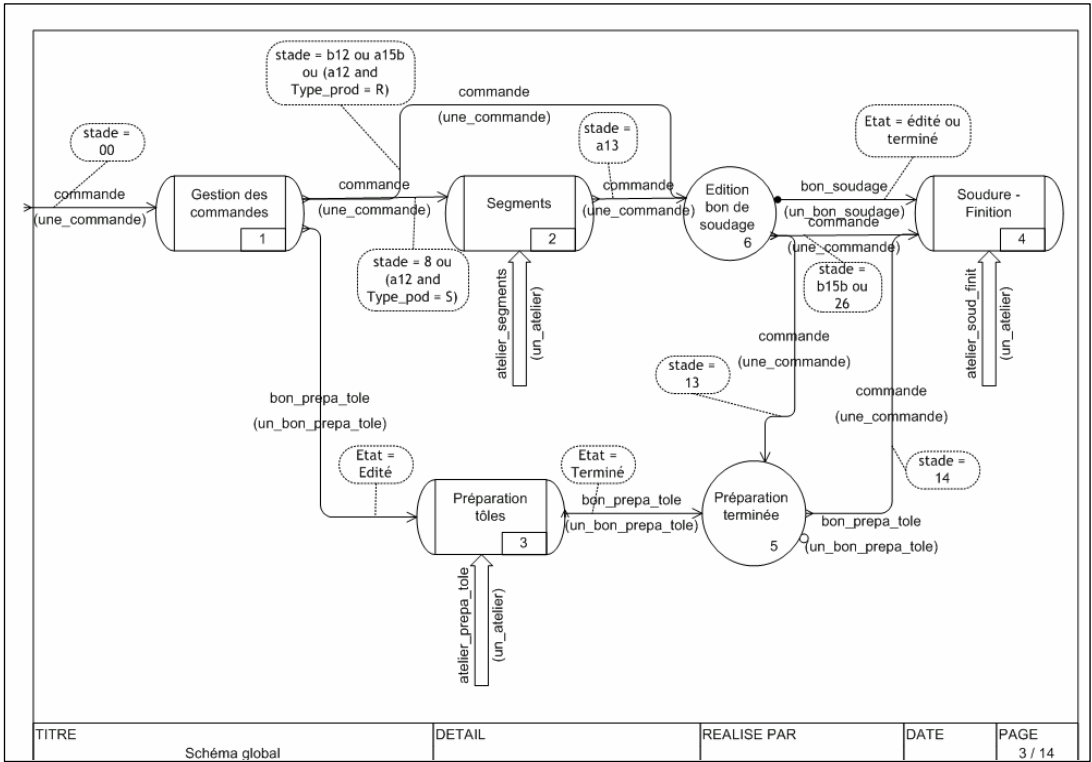


Рисунок 7 – Фрагмент общей схемы реального процесса в ALIX.

Заключение

Промышленное использование метода ALIX позволило выявить определенные недостатки. Например, трудности описания непрерывных процессов. В целом же реакция специалистов предприятий, познакомившихся с методом, была весьма обнадеживающей. Их мнение было позитивным, отмечалась полезность разработанных моделей для общего анализа происходящего (даже без имитации). Это подтвердило нашу уверенность в полезности простого графического языка как средства коммуникации между исследователями-консультантами и предприятиями. Несмотря на простоту, язык позволил нам описать все без исключения ситуации, встретившиеся в изучаемых процессах. С другой стороны, в языке не было выявлено ничего бесполезного.

Опыт применения высветил возможные пути дальнейшего совершенствования метода. Это его расширение на непрерывные (дискретно-непрерывные) процессы, добавление возможностей описания структуры предприятий и расширение возможностей описания методов и алгоритмов принятия решений.

Список литературы

- [1] Fox M. S., Gruninger M. Enterprise Modelling// The American Association for Artificial Intelligence.
- [2] Kalpic B., Bernus P. Business process modelling in industry-the powerful tool in enterprise management. Elsevier.
- [3] Vernadat, F. 'UEML : Towards a Unified Enterprise Modeling Language// Actes de la Conférence MOSIM'01, avril 2001, Troyes (France).
- [4] Iassinovski S., Artiba A., Bachelet V., Riane F. Integration of Simulation and Optimization for Solving Complex Decision Making Problems// Proceedings of the *International conference on industrial engineering and production management (IEPM'2001)*, Quebec City, Canada, August 20-23, 2000, 1186-1195 pp.
- [5] Ясиновский С.И., Артиба А., Башле В., Риан Ф. Интеграция имитационного моделирования и методов оптимизации для решения сложных проблем принятия решений // Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции / Под ред.: акад. В.П. Мясникова, акад. Н.А. Кузнецова, проф. В.А. Виттиха.- Самара: Самарский научный центр РАН, 2002.- 560 с., ISBN 5-93424-064-1. (Proceedings of the IV International Conference on Complex Systems: Control and Modeling Problems, June 17-23, 2002, Samara, Russia), pp. 266-272.
- [6] Vernadat, F. Techniques de modélisation en entreprise: applications aux processus opérationnels. Collection Gestion. Editions Economica.
- [7] Kim C.-H., Weston R. H., Hodgson A., Lee K-H. The complementary use of IDEF and UML modelling approaches. Computers in Industry. Elsevier.
- [8] Santarek K., Buseif I. M. Modelling and design of flexible manufacturing systems using SADT and Petri net tools// Journal of Materials Processing Technology 76. 212-218 pp.
- [9] V.Emelyanov, S.Iassinovski, T.Shtautmaster IDEF-RAO: Methods for industrial system functional structure analysis// Proceedings of Workshop on production planning and control. Mons, Belgium 9-11 September 1996 pp. 387-390.
- [10] Paige R. F., Ostroff J. S., Brooke P. J. Principles for modelling language design. Information and Software Technology. Elsevier.
- [11] Monteiro T. Conduite distribuée d'une coopération entre entreprises : le cas de la relation donneurs d'ordres-fournisseurs. Thèse du Laboratoire d'Automatique (LAG) de l'Institut National Polytechnique de Grenoble.
- [12] Talbi A. Contribution à l'intégration des fonctions de l'entreprise. Application aux fonctions production et maintenance. Thèse de l'Université Mohammed 5 à Rabat (Maroc).

ИНТЕГРАЦИЯ МОДЕЛЕЙ ЗНАНИЙ НА ОСНОВЕ ОБЪЕКТНО-КОГНИТИВНОГО АНАЛИЗА

М.Б. Гузаиров, Р.А. Бадамшин, Б.Г. Ильясов, Л.Р. Черняховская, Р.А. Шкундина

Уфимский государственный авиационный технический университет

450000, Уфа, ул. К.Маркса,12, Россия

lr_chern@webmail.aport.ru

тел: +7 (3472) 233-128. Факс (83472) 234-778

Ключевые слова: база знаний, модель знаний, объектно-когнитивный анализ, правило, прецедент, водная экосистема

Abstract

Gathering and integration of knowledge derived from various sources has been an important technique of decision-making in complex object management process. The approach for the knowledge model integration based on the Object Cognitive Analysis is suggested. This paper proposes a hierarchical knowledge base structure, rules and cases models of knowledge presentation, and explores its applicability to the problem of ecological systems management. The decision making algorithm was evaluated on the fuzzy system. The experiments indicate that these process control techniques are adequate to the experts' notion of the management process being investigated.

Введение

Статья посвящена интеграции различных моделей знаний в ситуационной базе знаний на основе предложенного подхода к управлению в критических ситуациях (КС).

На первом этапе разработки базы знаний рекомендуется провести объектно-когнитивный анализ, в процессе которого происходит формирование понятий ПО и определение отношений между ними [1]. *Объектно-когнитивный анализ (ОКА)* предметной области интегрирует методы объектно-ориентированного анализа, онтологического анализа и семантического анализа процесса управления в КС.

Разработаны следующие принципы объектно-когнитивного анализа:

- 1) принцип иерархической декомпозиции знаний, основанный на применении различных форм абстрактного представления знаний и полиморфизме описания отношений между объектами семантической сети и механизмов логического вывода;
- 2) принцип интеграции онтологического анализа и семантического моделирования предметной области на основе иерархии понятий;
- 3) принцип введения понятий в онтологии предметной области на базе объектно-ориентированного моделирования процесса управления в критических ситуаций;
- 4) принцип наследования свойств, основанный на отношении «класс – подкласс» в описании понятий в различных формах абстракции представления знаний;
- 5) принцип определения меры сходства в семантических сетях на основе кластерного анализа описаний прецедентов КС в терминах предметной области и разработки таксономической структуры тезауруса.

Результатами объектно-когнитивного анализа предметной области являются описания парадигматических (ассоциативных, иерархических, функциональных и синонимических) отношений между абстрагированными понятиями и сущностями, являющимися базовыми объектами предметной области (когнитивными элементами семантической сети), в терминах предмет-

но-ориентированного тезауруса. Полученные описания являются основой разработки базы знаний информационной системы поддержки принятия решений.

1 Объектно-когнитивный анализ процесса управления экосистемой

Рассмотрим применение объектно-ориентированного моделирования для процесса управления экосистемой в КС.

В рассматриваемой предметной области система поддержки принятия решений может иметь слишком много релевантных деталей в рамках единственной абстракции, что затрудняет ее интеллектуальную управляемость. В подобных случаях можно обеспечить такую управляемость путем декомпозиции модели в иерархию абстракций. Иерархия позволяет управляемым образом вводить релевантные детали. Абстракции на любом заданном уровне иерархии позволяют (временно) игнорировать многие релевантные детали при понимании абстракций на следующем, более высоком уровне.

Множественное наследование отображается через сетевую рекурсию, которая задает сеть отношений между экземплярами родительской и дочерней сущностей. Это случай, когда сущность находится сама с собой в отношении «многие ко многим». На рисунке 1 показана модель «сущность - связь» (англ. Entity Relationship Diagram, ERD) в нотации Unified Modeling Language, показывающая иерархию классов представления знаний в базе знаний.

Одно из преимуществ такой "иерархии абстракций" заключается в возможности доступа различных пользователей к модели на разных уровнях абстракции. Таким образом, единственная модель может совместно использоваться несколькими различными пользователями без нарушения их требований к доступу. Другим достоинством иерархии абстракций является повышенная стабильность модели в процессе развития приложения или самой системы. Изменения в тех деталях, которые игнорируются на более высоких уровнях модели, не будут затрагивать эти уровни. Некоторые изменения, конечно, будут при этом распространяться с нижнего уровня на верхний. В соответствии с этой техникой объекты, относящиеся к предметной области, классифицируются на некоторое число категорий или классов на основании их общих свойств.

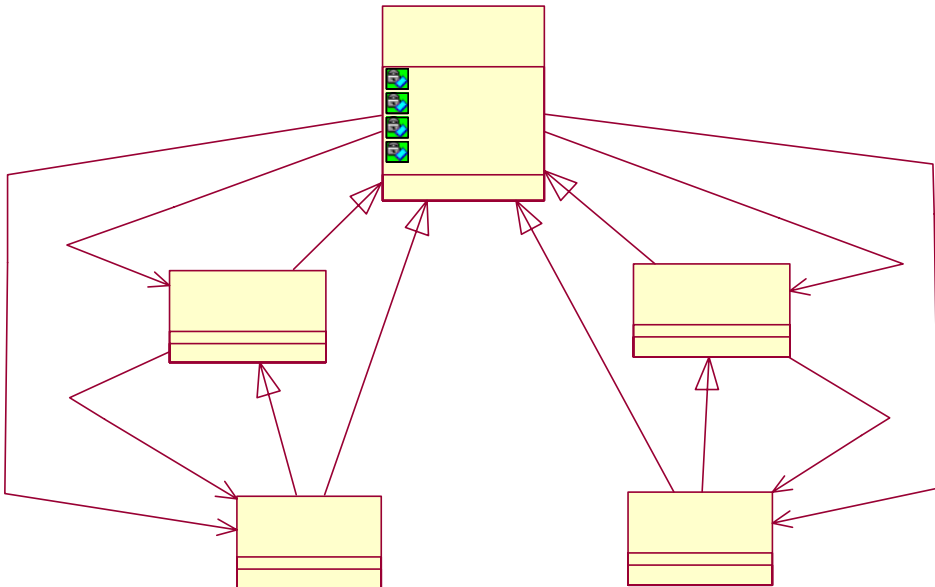


Рисунок 1 – Модель сущность-связь

В структуре базы знаний реализован метод наследования свойств, определяющий отношения в иерархии классов «Правило – Категория прецедентов - Прецедент». Приведем пример отношений обобщения между классами объектов в конкретной предметной области.

На рисунке 2 показана схема структуры базы правил и прецедентов, предназначенной для управления водной экосистемой, в том числе и в критических ситуациях. Выделенные классы правил образуют иерархию, соответствующую установленной иерархии объектов модели управления водной экосистемой. На приведенной модели видно, что в классах нижних уровней иерархии наследуются свойства классов, определенные в классах верхнего, “родительского” уровня (понятия, атрибуты и поведение классов объектов предметной области) в результате объектного моделирования.

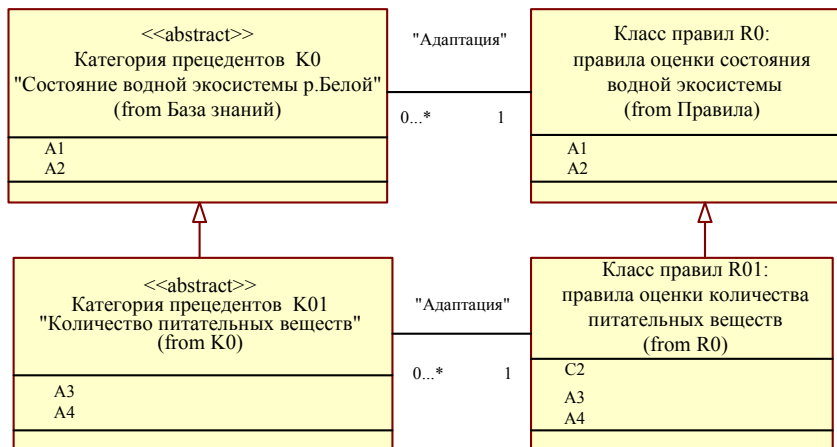


Рисунок 2 – Интеграция различных моделей представления знаний на основе наследования свойств

2 Разработка иерархической структуры базы знаний по результатам объектно-когнитивного анализа

В сложных системах, состояние которых характеризуется большим количеством признаков, возникает проблема обеспечения полноты базы знаний [2]. Известно, что количество правил, необходимых для описания n входных переменных, значение которых определяется с использованием m термов, равно $R=mn$.

Следовательно, количество правил увеличивается экспоненциально с увеличением количества входных переменных, что приводит к увеличению времени логического вывода решения. База знаний, содержащая большое количество правил, сложна в восприятии, редактировании и использовании. Но процесс принятия решений требует, тем не менее, полного осмысления критической ситуации. Это противоречие разрешается путем построения иерархической базы знаний, таким образом, что класс правил может быть классом верхнего уровня или подклассом, ассоциированным с отдельным объектом предметной области (см. рисунок 3).

Декомпозиция базы правил производится в соответствии с иерархией объектов, установленной в результате объектного моделирования. Таким образом, каждое правило включает ограниченное количество входных переменных. Выход модуля базы знаний i -го уровня является входом модуля базы знаний $(i+1)$ уровня.

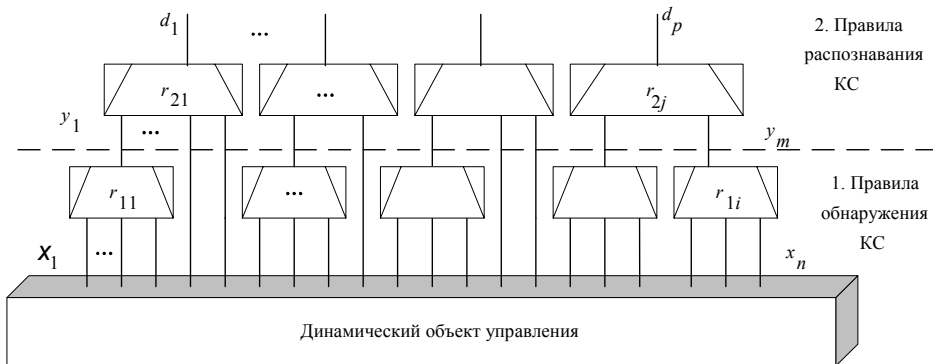


Рисунок 3 – Иерархия правил принятия решений

Задачей правил первого уровня является оценка нарушения границы области допустимых значений по отдельным переменным состояния, наиболее полно характеризующим классы КС, так как, по определению каждый класс КС характеризуется нарушением предельно допустимых значений одной или нескольких переменных состояния.

Правила первого уровня формируют решения об установлении нарушения границ области допустимых значений критических переменных в отдельной подсистеме сложного динамического объекта (например, водной экосистемы). Второй уровень объединяет сигналы с выходов правил первого уровня в вектор, характеризующий состояние работоспособности объекта в данный момент. Следовательно, пространство признаков классификации КС $Y(t)$, являющихся входами правил второго уровня, является вырожденным отображением пространства признаков $X(t)$. Правило второго уровня координирует решения правил первого уровня, производя объединение информации с целью более точного распознавания классов КС. Необходимость такой координации обусловлена взаимозависимостью критических состояний отдельных подсистем.

Назначением правил второго уровня является распознавание класса критической ситуации. Пример правила распознавания КС:

$$r_{2e}: \text{ЕСЛИ } y_r=B_r \text{ И } \dots \text{ И } Y_s=B_s,$$

$$\text{ТО } z^\gamma = h,$$

где z^γ – класс КС S_r , h – имя класса КС из множества наименований W^h .

Приведенные правила распознавания КС обеспечивают более точную аппроксимацию границы области допустимых состояний по сравнению с существующими логическими схемами сигнализации, которые обозначают границу области критических режимов по какому-нибудь одному параметру, одной точкой.

Правила принятия решений 2-го уровня определяют также типовые решения для распознанных классов КС:

$$r_{2k}: \text{ЕСЛИ } z^\gamma = h \\ \text{ТО } d_k=D^h,$$

где d_k – типовое управляющее решение значения D^h .

Правила уровня 3 определяют значения обобщенных индикаторов состояния объекта:

$$r_{2k}: \text{ЕСЛИ } z^\gamma = h \\ \text{ТО } a_f=1,$$

где a_f – признак нахождения индикатора состояния объекта в области критических режимов.

Разработка иерархической нечеткой базы правил включает такие специфические этапы, как декомпозиция правил принятия решений в соответствии с результатами кластерного анализа входных переменных и оценка непротиворечивости иерархии правил в соответствии с принципом “наследования свойств”.

3 Пример базы знаний для оценки состояния водной экосистемы

Приведем пример разработанной нечеткой базы знаний для оценки возможности ухудшения качества воды, представленный на рисунке 4. Этот показатель зависит от целого ряда параметров (прозрачность “clarity”, жесткость “rigidity”, содержание азота “nitrogen”, биомасса “biomass” и др.). Нечеткая система содержит три блока правил, составляющих систему правил, характеризующую объект, и организованных в иерархию в соответствии с моделями структуры базы знаний. Также на рисунке 4 приведен фрагмент таблицы, содержащей нечеткие правила.

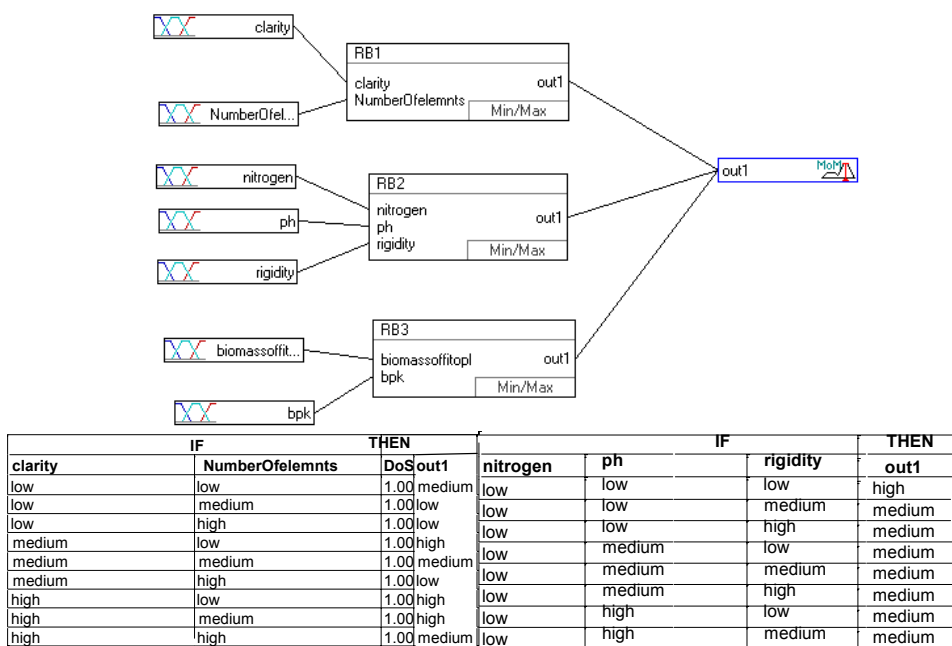


Рисунок 4 – Структура нечеткой базы знаний

Анализ адекватности приведенной нечеткой модели показал, что влияние изменения значений входных нечетких переменных на выходную переменную соответствует представлением экспертов о рассматриваемой критической ситуации.

Заключение

В статье представлен способ формирования решений на основе ситуационной базы знаний, включающей различные модели представления знаний: правила принятия решений, прецеденты и объектно-ориентированные модели сценариев управления. Способ интеграции различных моделей представления знаний базируется на принципах объектно-когнитивного анализа. Организация поддержки принятия решений в критических ситуациях, возникающих в экологической системе, на основе разработанной базы знаний показала, что результаты соответствуют представлениям экспертов о рассматриваемой предметной области. Реализация си-

туационной базы знаний с использованием теории нечетких множеств осуществлена с использованием инструментального средства FuzzyTech. Интеграция знаний, формализованных с использованием различных моделей, преследует цели обеспечения большей выразительности представления и большую эффективность обработки знаний. В дальнейшей работе предполагается усовершенствование алгоритма поиска решений в базе знаний в условиях неопределенности, с учетом различных априорных сведений о процессе управления экосистемой.

Список литературы

- [1] Кусимов С.Т., Черняховская Л.Р., Осипова И.В., Старцева Е.Б.. Формирование управляющих решений в критических ситуациях на основе инженерии знаний. Труды V международной конференции «Проблемы управления и моделирования в сложных системах» / Под ред. Академика В.П. Мясникова, академика Н.А. Кузнецова, профессора В.А. Витиха. Самара: Самарский научный центр РАН, 2003, с. 467 – 472.
- [2] Р.А. Бадамшин, Б.Г. Ильясов, Л.Р. Черняховская. Проблемы управления сложными динамическими объектами в критических ситуациях на основе знаний. – М.: Машиностроение, 2003 – 240с.

RDF-ТЕХНОЛОГИИ ПРЕДОСТАВЛЕНИЯ СЕРВИСОВ МОБИЛЬНЫМ ПОЛЬЗОВАТЕЛЯМ В БЕСПРОВОДНЫХ СЕТЯХ*

Н.И. Юсупова¹, Д.В. Попов¹, И.А. Вайнерман¹, Б. Кёних-Риз²

¹Уфимский государственный авиационный технический университет
450000, г. Уфа, ул. К.Маркса, 12,
yussupova@ugatu.ac.ru

Тел. (3472) 23-77-17. Факс (3472) 23-77-17

²Технический Университет Карлсруэ,
76128 Karlsruhe, Germany

koenig@ipd.uni-karlsruhe.de

Тел. +49/721/608-3968. Факс +49/721/608-7343

Ключевые слова: формирование запроса, передача данных, RDF, RDQL, HTML, XML, обработка запроса, онтология

Abstract

The questions of development and designing of the system for processing the queries from mobile users, storage, transfer and granting the access to the required information are considered in the article.

Введение

За последнее время на рынке мобильных устройств все большую популярность приобретают Personal Digital Assistant (PDA - Личный Цифровой Помощник) и smartphone (смартфон) – сотовые телефоны с расширенным набором функций, поддерживающие технологии Java2, Bluetooth и другие. Встает актуальный вопрос об использовании данного вида устройств при получении доступа к сетевой информации. Такие устройства в качестве достоинств имеют миниатюрные размеры, широкий круг предоставляемых функциональных возможностей, возможность интеграции с персональным компьютером (ПК) и подключения к глобальным и локальным сетям Internet/Intranet.

Использование указанных устройств порождает целый ряд проблем в сфере предоставления информации. Возникают такие задачи как: хранение и передача информации, обработка запросов от мобильных пользователей, предоставление пользователю информации в необходимом виде, представление данных на логическом и физическом уровнях, разработка механизмов обработки информации. В настоящее время существует разнообразные подходы к решению данного круга проблем. Наряду с применением традиционных баз данных (БД), базирующихся на реляционных подходах, приобретают все большую популярность такие технологии как веб-сервисы (web-services) и другие технологии, базирующиеся на использовании языков синтаксического и семантического описания данных. Получила большое распространение технология Java Beans, предоставляющая возможности гибкого использования удаленных процедур.

* Данное исследование частично поддержано грантом РФФИ 03-07-90242-в «Интернет-комплекс поддержки выполнения проектов фундаментальных исследований сложных систем с применением интеллектуальных технологий на базе экспертных систем» (2003-2005 гг.), а также Федеральной целевой программой "Интеграция науки и высшего образования Российской Федерации на 2002-2006 годы" (программное мероприятие 1.4 - "Привлечение иностранных партнеров к проведению совместных исследований и развитию интегрированных научно-образовательных структур", проект № П0039 "Фундаментальные исследования и новые технологии проектирования сложных технических систем").

Указанные подходы имеют свои достоинства и свои слабые стороны, но на данном этапе развития информационных технологий не существует пока унифицированных средств реализации механизма предоставления информации мобильным пользователям. Это обусловлено рядом факторов, таких как: специфичность и разнообразие информации, небольшие объемы памяти для хранения информации в мобильных устройствах, разнообразие платформ для реализации соответствующих механизмов.

Услуга или сервис – это предоставление необходимой информации пользователю в ответ на запрос, который он адресует системе. Запрос должен иметь форму адекватную (интуитивно понятную) мобильному пользователю, а ответ - вид, поддерживаемый мобильным устройством пользователя.

1 Логика работы системы

При реализации системы предоставления сервисов возникает ряд задач:

- 1) Разработка концепции построения системы, выделение отдельных функциональных составляющих системы и распределение ролей, которые они должны играть в процессе функционирования системы.
- 2) Проблема логического представления внутренних данных системы в базах данных. Существует много подходов к решению этой проблемы, начиная от традиционных уже реляционных таблиц или баз, где данные представлены в виде объектов, заканчивая разнообразными представлениями и описаниями данных с использованием языков разметки.
- 3) Проведение анализа всех составляющих системы, разработка алгоритмов функционирования отдельных частей, выбор технологий, которые будут применяться непосредственно при реализации каждой составляющей части. На этом этапе должны быть разработаны такие механизмы, как: построение и непосредственно обработка запроса от пользователя, механизм проведения запросов данных непосредственно в базах данных, механизм формирования информации, возвращаемой пользователю.

На рисунке 1 приведена схема логики работы системы и ее функциональные части. Можно выделить основные составляющие части системы: пользователь с мобильным устройством и три компоненты:

- *Треjder* - Удаленный сервер (база данных), предоставляющий пользователю информацию о том, в какой базе данных он может найти информацию по интересующей его теме.
- *Медиатор* – Интерфейсная часть системы, которая должна служить «буфером» при взаимодействии пользователя и непосредственно базы данных. Здесь производится первоначальная обработка запроса пользователя и генерирование данных в виде, отвечающем требованиям мобильного устройства пользователя.
- *Домен* - База данных для хранения информации, должна иметь также механизм для обработки и передачи информации.
- *Пользователь* – человек с каким-либо мобильным устройством, непосредственно использующий предоставляемые сервисы. Подразумевается то, что на мобильном устройстве установлена клиентская часть системы, имеющая удобный пользовательский интерфейс для формирования запросов.

Предлагается следующий механизм работы системы предоставления сервисов.

В соответствии с постановкой задачи, пользователь вводит необходимые параметры для запроса, и они пересылаются в медиатор, где из них формируется первоначальный запрос. В соответствии с данным описанием возвращаемая информация будет иметь тот или иной вид. К примеру, это может быть как просто текстовая информация, так и данные в HTML формате с картинками, предназначенные для просмотра на PDA или портативном компьютере.

В медиаторе эти параметры преобразуются в RDQL запрос. Далее запрос посылается в домен, где непосредственно выполняется. Полученные результаты вновь преобразуются к RDF

формату. После того как необходимые данные получены в базу данных медиатора, пользователю посылается предварительный ответ (Ответ 1 на рисунке1), содержащий информацию относительно данных, которые требуются пользователю. Далее, для получения более детального ответа от системы, пользователь вводит новые параметры, которые посылаются в медиатор. По новым параметрам в медиаторе производится выборка более приоритетной информации необходимой пользователю. Для этого формируется новый RDQL запрос. На рисунке 1 это представлено преобразованием Парам.2 -> RDQL2. Далее данный запрос запускается уже не в домене, а непосредственно в медиаторе. Полученный результат возвращается пользователю в форме, отвечающей техническим характеристикам мобильного устройства пользователя.

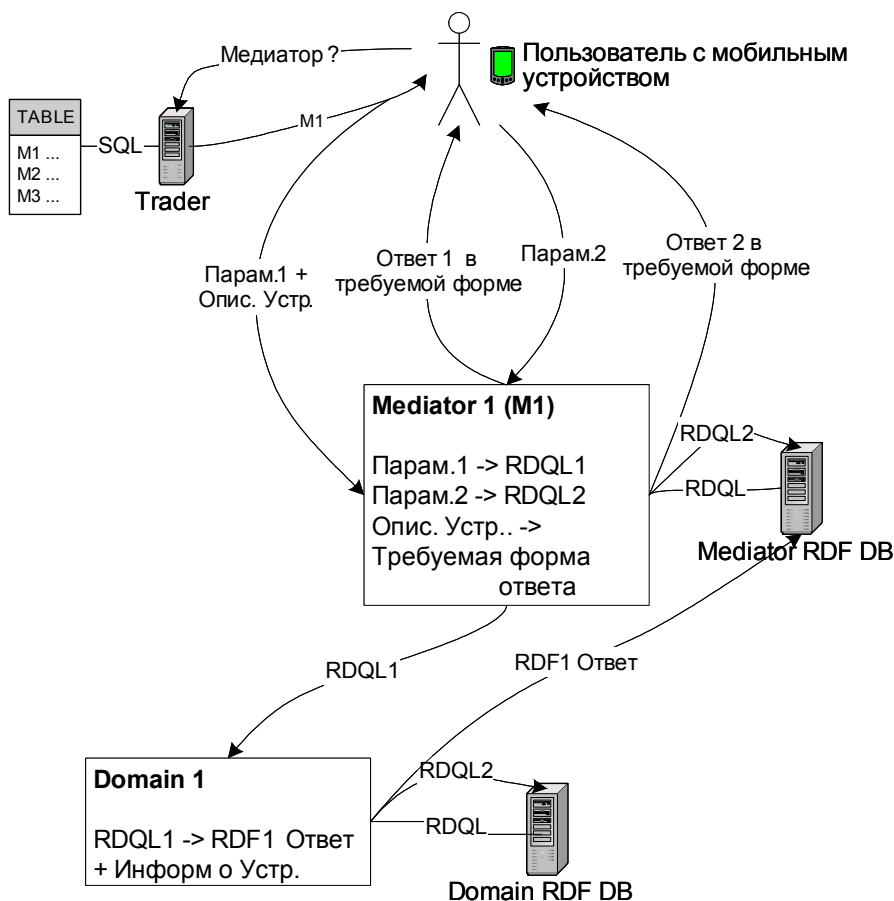


Рисунок 1 - Схема логики работы системы.

Для решения поставленных задач был проведен анализ современных средств хранения и обработки данных. Сравнение основных характеристик различных технологий представлено в таблице 1.

Таблица 1 - Основные характеристики средств хранения и обработки данных.

Характеристики	Реляционные БД	HTML	XML	RDF
Разметка данных	-	+	+	+
Язык запросов	+	-	+	+
Древовидное представление данных	-	-	+	+
Представление синтаксиса	-	-	+	+
Представление семантики	-	-	-	+
Расширяемость	+	-	+	+

2 Технология RDF

Для реализации прототипа системы была выбрана технология RDF. RDF (Resource Description Framework, язык описания ресурсов) [1] является разработкой WWW-консорциума (W3C). RDF является метаязыком, но наиболее важным отличием от других языков метаданных, таких как XML (eXtensible Markup Language), является то, что при помощи XML можно описать данные на синтаксическом уровне, а при использовании RDF информацию можно описать на семантическом уровне. Этот подход описания данных основан на принципе онтологий. Онтология это один из методов представления зависимостей между данными. Информация представленная в виде ресурсов описывается при помощи предикатов определяющих отношения между различными ее частями.

Язык описания ресурсов применяется в рассматриваемой системе как основное средство представления информации. Для практического применения данного средства необходимо использование схем. Схема – это также документ RDF, в котором содержатся спецификации для построения других RDF документов.

В прототипе системы применяются данные, представленные в RDF нотации. Для их обработки и выполнения запросов необходимо использовать язык запросов для RDF данных.

RDQL (RDF Query Language) [4] - язык реализации запросов для данных представленных в RDF формате. По синтаксису язык напоминает SQL, но имеет немного другую структуру построения запросов. Запрос имеет следующий вид:

```
SELECT ?x
WHERE (?x, <name>, "John")
```

В прототипе системы в качестве сервисов пользователю предоставляется информация о мобильных устройствах. В роли мобильных устройств могут быть сотовые телефоны, PDA, портативные компьютеры или любые мобильные информационные устройства. Для прототипа системы используется разработанная RDF схема. При использовании данной схемы можно в формате RDF описать любое мобильное устройство.

В RDF-схеме [2] используются такие типы данных RDF как класс, свойство, подкласс. Информацию относительно RDF, его типов данных и схем приведена в [7]. При помощи схемы некоторое мобильное устройство можно представить в следующем виде: (рассмотрим на примере сотового телефона Siemens ME45)

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:md="http://localhost/Mob_dev/mob_dev_sch.rdf#">
<md:Mobile_device rdf:ID="SiemensME45">
  <md:Input_unit >
```

```

    <md:Custom rdf:ID="Mobile_phone_keypad"/>
  </md:Input_unit >
  <md:Audio>
    <md:Beep rdf:ID="Mobile_phone_speaker" />
  </md:Audio>
  ...

```

В данном примере используется пространство имен md, указывающее на RDF схему mob_dev_sch.rdf. Эта схема уже упоминалась выше. Объект класса Mobile_device (Мобильное устройство) имеет несколько свойств, первое свойство это устройство ввода Input_unit (Устройство ввода), которое в качестве своего значения имеет объект класса Custom (Специальное), представляющий обычную клавиатуру мобильного телефона. Следующее свойство это аудио устройство Audio (Аудио устройство), которое имеет значение в виде объекта класса Beep (Низкое качество), представляющего обычный динамик мобильного телефона.

Для представления данных в RDF формате необходима соответствующая схема, в которой в объектно-ориентированных терминах должны быть описаны все логические понятия, связанные с той областью знаний, которая должна быть представлена в системе. Схема для такой области, как мобильные устройства, должна содержать понятия, при помощи которых можно описать практически любое мобильное устройство. В качестве таких понятий могут выступать такие части мобильного устройства как: устройство ввода, устройство вывода, аудио устройство, коммуникационное устройство. Графическое представление разработанной схемы приведено в [10].

Непосредственное описание некоторого мобильного устройства в RDF формате с использованием вышеописанной схемы можно представить в синтаксисе XML следующим образом [11]:

```

<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:md="http://localhost/Mob_dev/mob_dev_sch.rdf#">
  <md:Mobile_device rdf:ID="SiemensME45">
    <md:Input_unit >
      <md:Custom rdf:ID="Mobile_phone_keypad"/>
    </md:Input_unit >
    <md:Audio>
      <md:Beep rdf:ID="Mobile_phone_speaker" />
    </md:Audio>
  ...

```

Здесь xmlns:md="http://localhost/Mob_dev/mob_dev_sch.rdf#" - это пространство имен, относящееся к схеме mob_dev_sch.rdf, которая и является вышеприведенной схемой. В данном примере представлен непосредственно сотовый телефон SiemensME45, который является объектом класса Mobile_device (Мобильное устройство). У данного объекта есть свойства, описанные в соответствующей схеме, как свойства класса Mobile_device (Мобильное устройство). К примеру, имеется свойство <md:Input_unit >, в качестве значения которого выступает объект класса Custom: <md:Custom rdf:ID="Mobile_phone_keypad"/>, представляющий собой обычную клавиатуру сотового телефона. И так далее, по аналогии с рассмотренным свойством, описываются все остальные свойства, представленные в схеме mob_dev_sch.rdf.

Заключение

Проведен краткий анализ современных технологий хранения и обработки данных. Произведено проектирование и реализация прототипа системы предоставления услуг мобильным пользователям. Разработано RDF-представление данных для прототипа системы. Реализованы механизмы обработки данных и реализации запросов к данным в RDF-формате. Проведен анализ возможностей RDF технологии при хранении и обработке данных.

Список литературы

- [1] Resource Description Framework (RDF) Model and Syntax Specification W3C Recommendation 22 February 1999. <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
- [2] RDF Vocabulary Description Language 1.0: RDF Schema W3C Working Draft 23 January 2003. <http://www.w3.org/TR/rdf-schema/>
- [3] RDF/XML Syntax Specification (Revised). W3C Working Draft 23 January 2003. <http://www.w3.org/TR/rdf-syntax-grammar/>
- [4] Jena Tutorial for Release 1.4.0 HP labs. <http://www.hpl.hp.com/semweb/doc/tutorial/index.html>
- [5] Flanagan, David. Java in a nutshell : deutsche Ausgabe für Java 1.4 / David Flanagan. Dt. Übers. von Matthias Kalle Dalheimer. - 4. Aufl. - Beijing ; Köln : O'Reilly, 2003. - XVII, 1084 S.; (dt.) ISBN 3-89721-332-X Original: Java in a nutshell <dt.>
- [6] Flanagan, David. Java Examples in a Nutshell : der Beispielband zu Java in a Nutshell / David Flanagan. Übers.: Dorothea Reder. - 2. Aufl.- Beijing ; Köln : O'Reilly, 2001. - XIV, 648 S. : Ill.; (dt.) ISBN 3-89721-196-3 Original: Java examples in a nutshell <dt.>
- [7] www.w3.org. Официальный сайт WWW Consortium.
- [8] Niki Pissinou, Kia Makki, Birgitta König-Ries: A Middleware-Based Architecture to Support Mobile Users in Heterogeneous Environments. Proc. of RIDE 2000, San Diego, USA, February 2000
- [9] Koenig-Ries B., Popov D., Muelle J., Plechova O. "Multidimensional Query Result Navigation for Mobile Users" (186). In: Computer Science and Information Technologies CSIT'2002. University of Patras, Rion, Patras, Greece, 2002, 8 p. <http://www.lar.ee.upatras.gr/csit2002/>. Версия на CD-ROM. (Proceedings of the 4th International Workshop)
- [10] N.I. Yussupova, B. König-Ries, D.V. Popov, I.A. Vainerman Data Structures for a System to Dynamically Provide Mobile Users with Information in Wireless Networks // Proceedings of the 5th International Workshop on Computer Science and Information Technologies CSIT'2003. Ufa, Russia, 2003, Vol. 1, pp. 100-107.
- [11] Попов Д.В., Вайнерман В.А. Прототип системы предоставления сервисов мобильным пользователям в беспроводных сетях // Принятие решений в условиях неопределенности: Межвузовский научный сборник. – Уфа: УГАТУ, 2003. –С.14-22.

ОНТОЛОГИЧЕСКИЙ АНАЛИЗ ПЛАТЕЖНЫХ СИСТЕМ И МОДЕЛИ ИХ ПОВЕДЕНИЯ

А.М. Кистанов

ОАО КБ «Солидарность»,
443099, Самара, ул. Куйбышева, 90
plast@solid.ru
тел: (8462) 70-44-49, факс: (8462) 70-89-50

Ключевые слова: онтологический анализ, платежная система, структура платежной системы, однородные и разнородные системы, транзакционные системы.

Abstract

The article is dedicated to analysis of payment systems as being entire units. Their behavioral models and interaction with other transaction systems are reviewed. Based on the analysis suggestions in regards to possible ways of payments systems development in the near future are made.

Введение

В настоящее время в мире существуют тысячи различных платежных систем. Об их количестве в России можно судить по справочнику ЦБ РФ, в который на начало 2004 года занесено около 200 различных систем. Элементами платежных систем являются сотни миллионов людей, миллионы технических устройств (банкоматы, терминалы), десятки тысяч кредитных организаций и т.д.

Если классифицировать платежные системы по территориальному охвату, то их можно подразделить на международные системы, действующие на территории многих государств (VISA, Master Card, JCB, AmEx), национальные системы, действующие преимущественно на территории одного государства (Золотая Корона, Юнион Кард, STB-Кард, Рапида, Comras+ в России; Geld-karte в Германии; Proton в Бельгии и т.д.), локальные системы, действующие на территории региона (НСС, Солидарность, действующие в Самарской области).

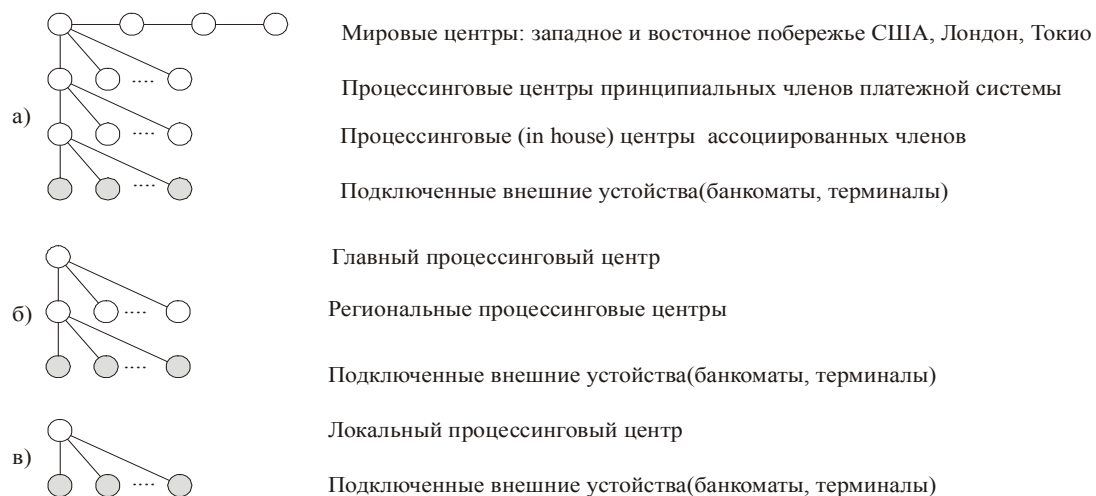


Рисунок 1 - Обобщенная структура а) глобальной, б) национальной, в) локальной платежных систем.

На рисунке 1 показана обобщенная структура систем с различным территориальным охватом. Если классифицировать системы по количеству эмитентов, то надо различать одноэмитентные и многоэмитентные системы; по процессинговым решениям это RS2, Open Way, Base24, ТРП и т.д.

В то же время, несмотря на обилие различных классификаций и возможностей, на практических конференциях, посвященных данной тематике [1], высказывалось мнение, что продвижение платежных систем основывается в большей степени на соображениях финансового и инженерного характера при недостаточном внимании к научным исследованиям в этой области. Вследствие этого совокупные затраты на внедрение систем в два, а порой и более раз, превышали требуемые. Чтобы по возможности восполнить этот пробел, проведем онтологический анализ [2] платежных систем и определим направление их развития на ближайшую перспективу.

1 Онтологический анализ платежных систем

В работе [2] представлена методика онтологического анализа сложных систем, в соответствии с которой рассматривается онтологический анализ первого и второго типа. Рассмотрим онтологию платежных систем России на «бытийном» описательном уровне, с точки зрения эволюции их структуры. Согласно [3] структура выражает то, что остается устойчивым, (относительно) неизменным при всех преобразованиях системы... Структура выступает интегрирующим фактором системы и детерминирует ее качество.

История платежных систем России насчитывает около пятнадцати лет и совпадает с началом формирования банковской системы страны. Примерно в одно время зародились российские платежные системы UNION CARD (Автобанк, Инкомбанк), STB CARD (СБС-АГРО), Золотая корона (Сибирский торговый банк), Optimum Card (Optimum bank), Сберкард (Сбербанк России). Появилось множество фирм, которые поставляли платежные решения на российский рынок: Скантек, Асофт, БПЦ, BGS. Любой банк, приобретая соответствующую технологию, мог создать на своей базе платежную систему. Часть систем в последующем прекратила свое существование вместе с закрытием расчетных банков (Optimum card). Другие продолжали работать. Как правило, это были системы, имевшие развитую многоуровневую структуру.

После того как основные игроки на рынке платежных систем определились, дальнейшие события показали, что локальные платежные системы стали активно проявлять интерес к интеграции. И если используемая системами технология позволяла это сделать, то они объединялись. Показательный пример - платежная межбанковская системы РАПИДА, которая объединила платежные системы, созданные по технологии фирмы Скантек (Москва). Так был решен вопрос о территориальном расширении системы, поскольку разрозненные точки отдельных участников стали общедоступными для всех. Система разрослась до общероссийской. Одной из основных проблем, которую при этом необходимо было решить, была обеспечение безопасности элементов системы. Еще раз подчеркнем, что такое объединение достигнуто только в рамках одной технологии. Объединения в рамках разных чиповых технологий пока не известны.

Чтобы система не пришла в запущенное состояние, чтобы в ней сохранялась, поддерживалась и даже развивалась структура, она должна быть открыта, то есть обмениваться с внешней средой массой и энергией [6].

Работающая платежная система постоянно требует увеличивающегося потока транзакций. Добиться этого можно двумя способами: увеличивать территориальное присутствие системы и расширять сервис, предоставляемый системой. Если этого не делать, то инициативу перехватят конкурирующие системы, с целью замкнуть весь поток транзакций на себя.

Параллельно с платежными системами развивались системы мобильной связи. Одни из них создавались как региональные (СМАРТС), другие как общероссийские (Мегафон, МТС) с

автоматическим роумингом на всей территории страны. Цели, которые преследуют сотовые операторы, так же предельно просты, как и у эмитентов карточек – привлечь максимальное количество клиентов с целью максимального получения прибыли. Что возможно через расширение сети обслуживания и качества предоставляемых услуг.

Следуя общесистемным закономерностям, платежные системы и системы мобильной связи с недавнего времени стали проявлять склонность к объединению. Но если раньше объединение касалось однородных (с точки зрения предметной области) систем, то теперь взаимодействуют разнородные системы.

Однако глобального объединения сети, как в случае с однородными системами, не происходит. Предположим, что платежная система A , имеющая a точек обслуживания, объединилась с платежной системой B , имеющей b точек обслуживания, как это показано на рисунке 2а. В этом случае в распоряжении клиентов системы A , как и в распоряжении клиентов системы B , имеется $a+b$ точек обслуживания. Если система мобильной связи K , имеющая k точек приема платежей, договорилась с платежной системой A о приеме платежей в a пунктах платежной системы, то у клиентов системы B , объединенной с системой A , такой возможности не будет. Это связано прежде всего с тем, что существующие финансовые протоколы ISO 8583, NDC, Diebold912 не позволяют в полной мере обеспечить подобное объединение.

Пример с сотовыми операторами далеко не единственный. Схожим образом развиваются системы обязательного медицинского страхования, где в качестве эмитентов выступают территориальные фонды, а в качестве точек обслуживания территориальные медицинские объединения.

Подведем краткие итоги:

- В одно время на одной территории может действовать множество платежных систем, направленных на привлечение клиентов, количество которых ограничено.
- Для того, чтобы сохранить свою жизнеспособность, платежные системы должны постоянно расширять свою структуру в пространстве (территориально) и во времени (количество и качество одновременно предоставляемых услуг в каждой точке).
- Наиболее естественно объединяются системы, созданные в рамках одной технологии.
- Системы, принадлежащие различным предметным областям, также объединяются между собой, образуя разнородные транзакционные системы. Однако такое объединение не транзитивно и не распространяется целиком на вновь созданную разнородную систему.

2 Модели поведения платежных систем

По количеству элементов, взаимосвязей между ними, интегративным свойствам платежные системы можно отнести к сложным социотехническим системам.

Поведение большинства элементов платежной системы носит стохастический характер. В [4] отмечается, что неопределенность формирует социотехническую систему как целостность. С точки зрения макроэкономики стохастическое поведение составляющих элементов преобразуется в поведение системы как целого [5]. Рассматриваемые далее модели отражают наиболее существенные [7] элементы поведения, необходимые для оценки поведения платежных систем в дальнейшем.

Определим функцию, цель и поведение платежной системы, исходя из методологии, представленной в [2].

Функция платежной системы - обеспечение финансовых потребностей и взаиморасчетов между участниками (элементами) системы.

Цель платежной системы – получение максимальной прибыли. Чем больше участников взаиморасчетов, тем больше прибыль, получаемая системой.

Поведение платежной системы, направленное на достижение цели, характеризуется расширением ее территориального присутствия, во-первых, и расширением объема и номенклату-

ры предоставляемых услуг, во-вторых. Это поведение обуславливается как внешним воздействием со стороны других платежных систем, имеющих сходные цели и поведение, так и внутренним воздействием со стороны многочисленных держателей пластиковых карт, желающих получать больше услуг в любое время и в любом месте.

В замкнутой системе, в которой количество клиентов достигло насыщения [8], сетевые свойства проявляются в том, что добавление любого внешнего устройства ведет к оттоку клиентов от действующих точек обслуживания. При равенстве доходов и расходов система достигает своего устойчивого (а точнее сказать «застойного») состояния, и дальнейшее ее территориальное развитие становится экономически невыгодным. Одно из условий равновесного состояния, полученного на основе модели [9], выражается формулой $b_{opt} = \sqrt{abc} / c$, где b_{opt} – оптимальное количество банкоматов в данной закрытой системе, a – совокупная комиссия, получаемая от всех операций в месяц, b – действующее количество банкоматов, c – совокупные затраты на эксплуатацию одного банкомата.

При наличии внешних воздействий, таких как инфляция, приход на рынок новых игроков, естественное желание персонала платежной системы иметь большую зарплату и т.д., платежная система как целостный элемент начинает искать новые пути получения прибыли: расширять географию присутствия системы и предоставляемый сервис.

В первом случае платежная система начинает договариваться с себе подобными об объединении сети обслуживания. Если экономические и технологические условия это позволяют, то такое объединение происходит.

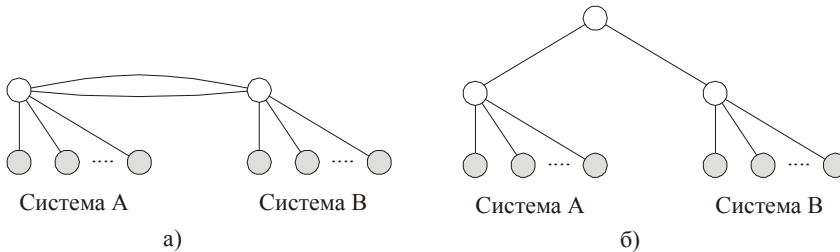


Рисунок 2 – а) горизонтальное и б) вертикальное объединение платежных систем.

На рисунке 2а показано вертикальное объединение, например объединение эквайринговой сети платежных систем Юнион Кард и СТБ в России. На территории Самарской области - это объединение сети платежной систем NCC и сети процессинговой компании РЕКАРД. Если объединяются три и более платежных систем, построенных по единой технологии, то, как правило, - это вертикальное объединение (рисунок 2б) с созданием главного процессингового центра, например Российская платежная система РАПИДА.

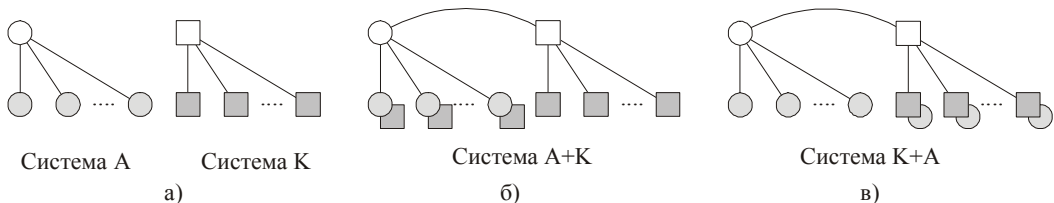


Рисунок 3 – а) две разнородные системы, б) объединение А+К двух разнородных систем, в) объединение К+А двух разнородных систем.

Расширение предоставляемого сервиса происходит как за счет внутренних возможностей системы, так и за счет объединения с другими транзакционными системами, имеющими отличные от платежных форматы транзакций.

На рисунке 3а показана платежная система *A* и, предположим, система мобильных платежей *K*. Следуя своей цели и внутреннему стремлению расширить предоставляемый сервис, система *A* включает в себя систему *K* (рисунок 3б). Подобное мы наблюдаем повсеместно, когда через банкоматы в режиме on-line оплачиваются услуги сотовых операторов. Возможно прямо противоположное объединение, когда система сотовой связи включает в себя платежную систему (рисунок 3в). Пример – Российская система мобильных платежей SimMP.

3 Модель поведения разнородных транзакционных систем

Устоявшегося определения – «транзакционная система», «транзакция», обобщающего эти термины для различных предметных отраслей, - не существует. Де-факто в сфере платежных технологий под транзакционными системами понимаются иерархичные системы, в которых межуровневые или межсистемные взаимодействия между элементами осуществляются при помощи точно описанных и стандартизованных протоколов. Любое событие, связанное с передачей информации в транзакционной системе, в соответствии с используемым протоколом, называется транзакцией.

Термин транзакция широко используется в клиент-серверной технологии при управлении СУБД ORACLE, MS SQL и других. С точки зрения свойств, так называемых АСИД (атомарность, согласованность, изоляция, долговечность), эти понятия совпадают. Однако с точки зрения структуры формата транзакций, многоуровневого использования - эти понятия различны. Исходя из предложенного определения, к транзакционными системам относятся платежные системы, системы расчетно-кассовых центров, системы мобильной связи. К ним можно отнести системы обязательного медицинского страхования и т.д.

Объединение двух разнородных систем не обладает свойством транзитивности, в первую очередь по причинам технологического характера, а еще точнее, из-за отсутствия в протоколах обмена соответствующих процедур. Простой пример. По карте международной платежной системы клиент сможет обслужиться в любой точке земного шара, получив деньги, товар или услугу. По той же карте клиент осуществит платеж в режиме on-line за мобильный телефон (оплатить за квартиру, обучение и т.д.) через банкомат того банка, который эмитировал карту. Однако он не сможет этого сделать абсолютно во всех банкоматах, где получает наличные. Для решения этой задачи следует изменить протоколы взаимодействия в системе, позволяющие передавать необходимое количество индивидуальных параметров субъекта. Это долгая, сложная, кропотливая работа. Однако, если этого не сделать, платежные системы потеряют часть своего бизнеса, клиентов, которые перейдут в руки других транзакционных систем, что в некоторой степени сейчас и происходит.

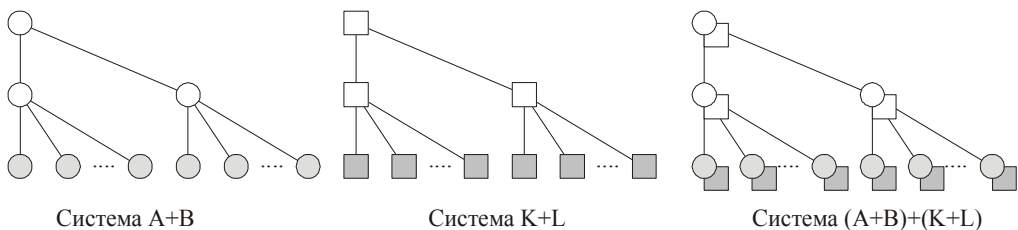


Рисунок 4 - Объединение многоуровневых разнородных транзакционных систем.

Кажущаяся очевидной изображенная на рисунке 4 структура системы $(A+B)+(K+L)$ в абсолютном большинстве случаев на практике неосуществима. Это связано прежде всего с тем,

что операция объединения систем не обладает транзитивностью, которая определяется форматом транзакций, принятых в системах. В лучшем случае форматы одинаковы для однородных транзакционных систем, и абсолютно разные для разнородных систем. Выход - структурный формат транзакций [10].

Заключение

В настоящее время в профессиональной среде, занимающейся развитием платежных систем и миграций на EMV стандарт, нет однозначного мнения относительно управления мульти-аппликационными возможностями, предоставляемыми этим стандартом. Однако, по крайней мере, есть два диаметральных мнения. Согласно первому - управление множественными приложениями на базе платежных должны осуществлять кредитные организации с учетом имеющегося колоссального опыта. Согласно второму мнению, эту функцию должны взять на себя операторы приложений, а банковское приложение будет одной из составляющих. Такими операторами приложений могли бы стать органы государственной власти, которые будут выдавать электронные удостоверения личности, планируемые в рамках программы «Электронная Россия». Точка зрения автора по данному вопросу следующая. Исходя из онтологического анализа платежных систем, наибольшей живучестью обладают платежные системы, имеющие транзитивные отношения с другими системами. То же самое можно сказать и о транзакционных системах иного рода. Поведение транзакционных систем в целом, и платежных систем в частности, характеризуется стремлением к объединению (равноправному или поглощающему). В выигрыше останутся те системы, которые сумеют создать или использовать наиболее приспособленные и стандартизированные протоколы межсистемного обмена.

Список литературы

- [1] Шестой Московский Международный Форум по платежным картам в России и Восьмая Международная конференция и выставка «Интеллектуальные карты России – 2003». Сборник материалов. (Москва, 24-26 ноября 2003 г.)
- [2] Виттих В.А. Онтологический анализ и синтез при управлении сложными открытыми системами // Проблемы управления и моделирования в сложных системах: Труды V международной конференции. – Самара: Самарский научный центр РАН, 2003. С.50-59.
- [3] Быстров А. Динамика структурности – опыт классификации. - [Http://www.n-torg/tp/ns/ds.htm](http://www.n-torg/tp/ns/ds.htm)
- [4] Виттих В.А. Целостность сложных систем // Проблемы управления и моделирования в сложных системах: Труды IV международной конференции. – Самара: Самарский научный центр РАН, 2002. С.48-58.
- [5] Попков Ю.С. Теория макросистем (равновесные модели). - М.: Эдиториал УРСС, 1999. – 320 с.
- [6] Жилин Д.М. Теория систем: опыт построения курса. - М.: Эдиториал УРСС, 2003. – 184 с.
- [7] Соколов Б.В., Юсупов Р.М.. Квалиметрия моделей – проблемы и пути развития // Проблемы управления и моделирования в сложных системах: Труды V международной конференции. – Самара: Самарский научный центр РАН, 2003. С. 22-37.
- [8] Романова Л., Бабаев А., Сафронов П. Эпидемия, или моделирование динамики эмиссии пластиковых карточек // Мир карточек. 2001. 7-8.
- [9] Кистанов А.М., Страдымов С.А. Сколько покупать банкоматов? / Платежи, системы, карточки. 2001. 10.
- [10] Кистанов А.М. Структурные методы построения форматов транзакций в развивающихся транзакционных системах // Вестник СамГТУ. Выпуск 14. Серия «Технические науки». – Самара: СамГТУ, 2002.

КОНЦЕПЦИЯ ИНТЕЛЛЕКТУАЛЬНОГО ИНТЕРНЕТ-ПОРТАЛА ЗНАНИЙ ДЛЯ ДОСТУПА К ИНФОРМАЦИОННЫМ РЕСУРСАМ ПО АРХЕОЛОГИИ И ЭТНОГРАФИИ¹

О.И. Боровикова¹, С.В. Булгаков², Ю.А. Загорюлько², Е.В. Сидорова¹, Ю.П. Холушкин³

¹Российский НИИ Искусственного интеллекта
630090, Новосибирск, пр. Лаврентьева, 6
olesya@iis.nsk.su
тел: +7 (3832) 32-83-59, факс: +7 (3832) 32-83-59

²Институт систем информатики СО РАН
630090, Новосибирск, пр. Лаврентьева, 6
zagor@iis.nsk.su
тел: +7 (3832) 34-29-91

³Институт археологии и этнографии СО РАН
630090, Новосибирск, пр. Лаврентьева, 17
hol@mail.archaeology.nsc.ru
тел: +7 (3832) 30-27-33

Ключевые слова: портал знаний, информационные ресурсы, археология, онтология науки

Abstract

The paper concerns a conception of specialized Internet-portal which provides access to information resources relating to archaeology and ethnography. An ontology which serves as a informational basis for portal supports integration of these resources in common information space and provides access to them through Internet. In order to user has a comprehensive notion about problem domain and integrated resources, ontology of knowledge portal integrates three relatively independent ontologies: ontology of science, ontology of scientific knowledge and subject domain ontology. Ontology of science describes concepts relating to organization of scientific activity. Ontology of scientific knowledge contains descriptions of various divisions of science, methods, objects and results of research. Subject domain ontology presents concepts of given branch of science, i.e. archaeology and ethnography. A portal also includes subsystem for collection of information about relevant Internet resources. This information is accumulated in database of portal and thus information space of portal is extended.

Введение

В настоящее время накоплен большой объем знаний и информационных ресурсов по гуманитарным наукам. Однако доступ к этим ресурсам весьма затруднен как из-за недостаточной формализованности этой области знаний, так и из-за несистематизированности самих ресурсов, их слабой структурированности и распределенности по различным Интернет-сайтам, электронным библиотекам и архивам.

Для решения задачи сведения ресурсов, относящихся к одной области знаний в единое адресное пространство, обеспечения возможности открытого удобного доступа к ним и поддержки их целостности нами предложена концепция специализированных Интернет-порталов знаний. Эта концепция положена в основу разрабатываемой нами технологии создания и сопровождения порталов знаний по любой научной тематике.

¹Работа выполняется при финансовой поддержке РГНФ (проект № 04-01-12045в), РФФИ (проект № 04-01-00884а) и СО РАН (Междисциплинарный интеграционный проект № 149).

Информационную основу порталов знаний составляют онтологии [1], содержащие наряду с традиционным описанием предметной области (ПО) соотнесенное с ним описание структуры и типологии соответствующих хранилищ данных и сетевых ресурсов [2, 3]. Кроме того, использование в качестве основы портала знаний онтологии, являющейся его декларативным компонентом, делает систему знаний портала легко расширяемой и настраиваемой – в нее могут интегрироваться как новые знания (например, о новых направлениях той или иной гуманитарной науки), так и новые срезы информационных ресурсов.

В данной работе обсуждается основанный на предложенной выше концепции подход к разработке специализированного Интернет-портала, обеспечивающего содержательный доступ к систематизированным знаниям и информационным ресурсам гуманитарных наук, в частности, к такому их подразделу, как археология и этнография.

1 Назначение и основные функции портала знаний

Портал знаний представляет собой специализированную информационную систему, снабженную эргономичным пользовательским web-интерфейсом.

С точки зрения пользователя, портал является тематическим Интернет-ресурсом, обеспечивающим возможность поиска и просмотра информации в рамках заданной предметной области (археология и этнография).

Как информационный ресурс портал:

- обеспечивает доступ к информации по различным аспектам и участникам научной деятельности, таким как: составляющие научной дисциплины (подразделы дисциплины, методы и техники исследования, используемые термины и понятия), персоналии исследователей, информация по группам, сообществам, организациям, включенным в процесс исследования;
- позволяет интегрировать близкие по тематике ресурсы, представленные в Интернет;
- предоставляет средства поиска интересующей пользователя информации в рамках всего информационного пространства портала;
- обеспечивает информационную поддержку пользователей ресурса (например, анонсирование разного рода событий и мероприятий);
- поддерживает гибкий пользовательский интерфейс, позволяющий учитывать предпочтения пользователя по работе с ресурсом и предоставляемыми сервисами.

2 Архитектура портала

При разработке архитектуры портала учитывались такие требования со стороны пользователей портала, как его полнота с профессиональной точки зрения, наглядность и простота использования.

Рассмотрим основные компоненты и модули портала.

База знаний объединяет тезаурус и онтологию портала.

Внутренняя база данных предназначена для хранения всей локальной информации, в частности, описаний ресурсов.

Модуль интеграции служит для подключения новых информационных ресурсов (источников данных) и поддержки унифицированного доступа к ним. При подключении таких ресурсов устанавливается соответствие между онтологией портала и системой терминов ресурса.

В состав портала также входит подсистема извлечения знаний и данных из сети Интернет – *коллекционер онтологической информации о ресурсах*. Входящие в ее состав специальные информационные агенты осуществляют поиск и сбор необходимой информации, которая накапливается во внутренней базе данных портала.

Пользовательский интерфейс предоставляет удаленный доступ к внутренней базе данных и базе знаний портала, а также к информационным ресурсам, проиндексированным в процессе его функционирования.

Для настройки портала на конкретного пользователя или группу пользователей в его состав включена *информационная модель пользователя*. В частности, модель пользователя содержит тематические предпочтения, список дополнительно подключаемых/отключаемых ресурсов, способ визуализации страниц.

Подсистема поиска информации предоставляет пользователю возможность задания запроса не только по ключевым словам, но и в терминах предметной области.

Основными элементами поискового запроса, заданного в терминах ПО, являются:

- *Понятия*, являющиеся элементами онтологии.
- *Ограничения*, которым должны удовлетворять найденные данные. Ограничения могут быть заданы в виде поискового шаблона определенного вида и/или логическими выражениями над значениями атрибутов понятий ПО.

Сформулированный таким образом запрос представляется как фрагмент онтологии с дополнительными ограничениями. Этот запрос преобразуется в один или несколько запросов к внутренней базе данных портала и/или к подключенным к portalу внешним структурированным источникам данных (СИД).

Для обеспечения поиска в СИД во внутренней базе данных в унифицированном виде хранятся описания схем данных внешних источников. Связывание схемы данных СИД с онтологией портала выполняется экспертом-настройщиком.

Поиск в неструктурированных источниках данных основан на использовании содержательных индексов, хранящихся во внутренней базе данных. Индекс источника данных строится на основе онтологии либо автоматически (коллекционером онтологической информации о ресурсах), либо вручную (экспертом).

3 Модель информационного наполнения портала

Информационную основу портала составляет онтология, важным компонентом которой является описание понятий, используемых в археологии и этнографии как отрасли науки, и соотнесенное с ней описание соответствующих сетевых ресурсов.

3.1 Онтология портала

Для достаточно полного и целостного представления пользователя о проблемной области и интегрируемых ресурсах, онтология портала знаний объединяет следующие относительно независимые онтологии: 1) онтологию науки, 2) онтологию научного знания и 3) онтологию предметной области, описывающую археологию и этнографию.

Онтология науки основана на предложенной в [4] онтологии, служащей для описания информации о научно-исследовательских проектах, и является ее развитием. Она включает следующие классы понятий, относящиеся к организации научной деятельности:

- *Ученые*. К этому классу относятся понятия, связанные с субъектами научной деятельности: исследователями, сотрудниками и членами организаций, исторически-значимыми персонажами и другими людьми.
- *Организации*. Понятия этого класса описывают различные организации, научные сообщества и ассоциации, институты, исследовательские группы и другие объединения.
- *События*. К событиям относятся такие понятия, как собрания, семинары, конференции, исследовательские поездки и экспедиции.
- *Публикации*. Этот класс служит для описания различного рода публикаций и материалов, представленных в печатном или электронном форматах (монографии, статьи, отчеты, труды конференций, периодические издания, фото- и видеоматериалы и др.).

- *Деятельность.* В этот класс входят понятия, описывающие научно-организационную или научно-исследовательскую деятельность – проекты, программы и т.п.

Онтология научного знания содержит метапонятия, задающие структуры для описания рассматриваемой предметной области. К таким метапонятиям относятся: раздел науки, метод, объект и результат исследований.

Онтология предметной области описывает научную дисциплину в целом как раздел науки и включает формальное и неформальное описание понятий и отношений между ними. Онтология опирается на словарь-тезаурус естественно-языковых терминов, описывающих значимую лексику предметной области.

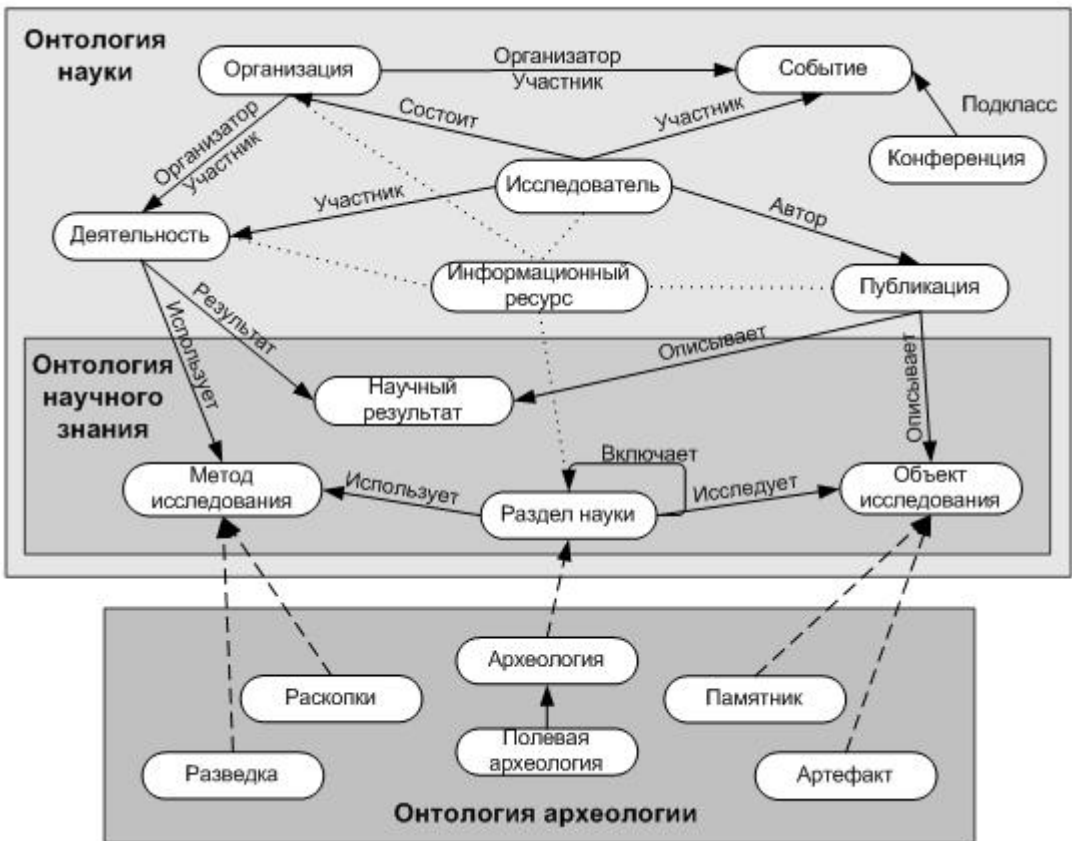


Рисунок 1 – Упрощенная схема онтологии портала.

На рисунке 1 представлен эскиз общей схемы онтологии портала. Он включает онтологии науки и научного знания, а также соотнесенную с ними онтологию археологии. На данной упрощенной схеме показаны не все связи, существующие между изображенными на ней понятиями. Так, не указаны связи, отвечающие за информационное наполнение ресурсов (информационный ресурс “описывает” событие), связи с объектами исследования (“кто открыл”, “в какой экспедиции открыт”), ассоциативные связи, не показана также иерархия разделов науки и научных направлений. Но в целом схема отражает основные понятия и связи между ними и является основой для построения полной модели.

Важным компонентом информационного наполнения портала является описание информационных ресурсов. Описание информационного Интернет-ресурса включает специфические

атрибуты и связи, определяющие его взаимоотношения с элементами онтологии. Набор атрибутов и связей основан на стандарте Dublin Core [5].

В информационном пространстве портала интегрируются следующие типы ресурсов:

- неструктурированные ресурсы – текстовое представление данных;
- слабоструктурированные ресурсы – например, html-документы;
- структурированные ресурсы – внешние базы данных, к которым есть права доступа.

3.2 Представление знаний по археологии и этнографии

Описание предметной области портала основывается на системной классификации археологической науки, предложенной Ю.П. Холюшкиным и Е.Д. Гражданниковым в [6] и развиваемой в настоящее время.

Системная классификация состоит из фрагментов определенной универсальной структуры. Стандартный классификационный фрагмент может быть представлен в виде семантической карты (см. рисунок 2), которая служит геометрической моделью фрагмента. Расположение элементов фрагмента определяется позиционной и ранговой координатами, соответствующим критериям первичности-вторичности, антиэнтропийности-энтропийности и общности-частности понятий. Каждое понятие может давать начало фрагменту более низкого яруса, для которого оно служит фоновым понятием, т.е. данный фрагмент охватывает площадку данного понятия, располагаясь под ней. Таким образом, геометрической моделью классификационной системы может служить трехмерное классификационное пространство, осями которого служат позиционная, ранговая и ярусная координаты.



Рисунок 2 – Фрагменты системной классификации, предложенной Ю.П. Холюшкиным и Е.Д. Гражданниковым

Внутри отдельного фрагмента существуют горизонтальные и вертикальные смысловые связи, делающие каждый классификационный фрагмент системой в том смысле, что это – целостное образование, содержащее информацию не только в отдельных элементах, но и в их

упорядоченных сочетаниях. При описании понятия на более низком ярусе возникают межярусные связи, устанавливающие отношения между элементами разных фрагментов.

Следует заметить, что подробная детализация разделов науки и понятий, имеющая место в рассмотренной выше классификации, и наличие большого количества типов связей между ними затрудняют использование данной классификации в полном объеме для навигации по информационному пространству портала. Так, например, некоторые популярные тематические разделы археологии являются достаточно специализированными и поэтому расположены в глубине классификационной иерархии, и выход на них требует от пользователя наличия большой профессиональной подготовки. Поэтому, для упрощения навигации по portalу используется заданная традиционным образом онтология, построенная на основе полной системной классификации археологической науки.

При разработке этой онтологии были выделены и представлены следующие аспекты классификации:

- основные направления археологии и этнографии – классификация по теоретическим разделам научной дисциплины и объектам исследования;
- классификация по временному признаку;
- классификация по географическому признаку;
- археологическая методология или научные подходы в археологии - классификация археологии по применяемым методам исследования.

Заключение

В докладе предложена концепция портала знаний, обеспечивающего эффективный и содержательный доступ широкому кругу пользователей к информационным ресурсам по археологии и этнографии.

Важным достоинством этого портала является то, что он обеспечивает доступ не только к собственным информационным ресурсам, но и поддерживает эффективную навигацию по релевантным ресурсам сети Интернет, проиндексированным в процессе его функционирования.

К настоящему времени разработана архитектура портала, онтологии науки и научного знания. Завершается разработка начальной версии онтологии археологии и этнографии и соответствующего ей словаря-гезауруса. Ближайшими задачами проекта являются:

- разработка web-интерфейсов пользователя и администратора системы;
- разработка коллекционера онтологической информации;
- автоматическое индексирование Интернет-ресурсов по археологии и этнографии.

Список литературы

- [1] Guariano N., Giaretta P. Ontologies and Knowledge Bases. Towards a Terminological Clarification // Towards Very Large Knowledge Bases. 1995-N.J.I.Mars (ed.) IOS Press, Amsterdam.
- [2] Боровикова О.И., Загоруйко Ю.А. Организация порталов знаний на основе онтологий. // Труды международного семинара Диалог'2002 «Компьютерная лингвистика и интеллектуальные технологии». Т. 2. -Протвино, 2002. С. 76-82.
- [3] Жигалов В.А., Загоруйко Ю.А., Нариньяни А.С., Россеева О.И. Предел однородности поиска в инернете // Системная информатика: Сборник научных трудов – Новосибирск: Наука, 2002. - Вып.8: Теория и методология программирования. С. 29-71.
- [4] Benjamins V.R., Fensel D. et. all. “Community is Knowledge! in KA2”, Proceedings of the KAW'98, Banff, Canada, 1998.
- [5] Using Dublin Core. <http://dublincore.org/documents/usageguide/>
- [6] Холушкин Ю.П., Гражданников Е.Д. Системная классификация археологической науки (элементарное введение в археологическое науковедение). Новосибирск: Изд-во ИДМИ Минобразования, Новосибирск, 2000. – 58 с.

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ В СИСТЕМЕ СОПРОВОЖДЕНИЯ КУРСА ЛЕЧЕНИЯ ЭЛЕМЕНТОЗОВ

Г.Б. Загорюлько

Российский НИИ Искусственного интеллекта
630090, Новосибирск, пр. Лаврентьева, 6
gal@iis.nsk.su
тел: +7 (3832) 32-83-59, факс: +7 (3832) 32-83-59

Ключевые слова: нутриенты, элементозы, коррекция микроэлементного баланса, моделирование предметной области, онтологический анализ, онтологический синтез

Abstract

The paper describes an approach to design of the system, which is intended for information support of a course of treatment of sickness concerned with microelement and vitamin imbalance of patient's organism. This system help to physicians to perform preliminary diagnostics of the imbalance, to pick up preparation and foodstuff, to institute a regimen and daily, weekly or monthly ration for patient. A subject domain model and operation scheme of the system is elaborated on the base of ontological analysis. The knowledge representation and processing facilities such as semantic networks, production rules and relational database is used for implementation of the system.

Введение

В любой области человеческой деятельности, где накоплены большие объемы знаний, очень актуальной становится задача структуризации и формализации этих знаний. Одной из таких областей, безусловно, является медицина вообще и, в частности, диагностика и лечение заболеваний, связанных с нарушением микроэлементного состава организма и авитаминозами. Конечно, существуют многотомные справочники, в которых описаны различные заболевания, синдромы и их многочисленные симптомы. Есть справочные пособия о витаминах, микро- и макроэлементах и их роли в организме человека. Есть перечни лекарственных препаратов и биологически активных добавок к пище (БАДов). Имеются обширные сведения о компонентах питания и составе продуктов. Однако имеющиеся знания разрознены, не всегда хорошо формализованы и структурированы, зачастую труднодоступны и неудобны для использования. Поэтому очень актуальной является задача создания для лечащего врача инструмента, который позволил бы ему всегда иметь под рукой необходимые справочные данные, помогал проводить диагностику, подбирать наиболее эффективные препараты с учетом побочных действий и имеющихся у пациента противопоказаний. Очень важно также не просто предоставлять пациенту список препаратов, а детально проработать и расписать схему их приема. Кроме того, необходимо уделить самое пристальное внимание диете. Практика показывает, что эта тема крайне редко затрагивается в кабинетах врачей, а ведь то, что мы едим, по мнению многих исследователей, является главной причиной наших недугов.

Целью настоящей работы является создание системы сопровождения курса лечения элементозов, которая является развитием описанной в работе [1] экспертной системы М-Корректор. Под элементозами понимается группа заболеваний, связанная с нарушением в организме человека баланса нутриентов (макро- и микроэлементов, витаминов, аминокислот). В работах [1, 2] отмечалась большая роль микроэлементов (МЭ) в поддержании здоровья. Рассматриваемая система работает на основе данных рентгенфлуоресцентного анализа волос, биохимического анализа крови, а также учета наблюдаемых у пациента симптомов.

1 Назначение и возможности системы

Рассматриваемая система предназначена для оказания помощи лечащему врачу на этапе лечения пациента. В ней собраны и систематизированы обширные знания и данные о микроэлементах и витаминах, о лекарственных средствах и продуктах питания, нормативы потребления элементов и их содержания в биосубстратах человека. Система обладает следующими возможностями:

- Ввод и редактирование данных о пациенте.
- Поиск и выдача справочной информации о роли микроэлементов в организме человека, нормах их содержания и потребления, о связанных с ними заболеваниях.
- Поиск и выдача справки о нутриентном наполнении, как отдельных продуктов, так и набора продуктов или определенных блюд. Т.е, система может вычислить, сколько и каких микроэлементов, витаминов вы съели, скажем, за обедом, сколько при этом вы получили калорий и холестерина, какую долю составили жиры, белки и углеводы, если вы введете список и количество продуктов, из которых был приготовлен ваш обед.
- Поиск продуктов и препаратов по заданным характеристикам. Например, определить список продуктов, содержащих определенное количество элементов и витаминов одной группы, не содержащих элементов другой группы и не превышающих в совокупности определенный уровень энергетической ценности.
- Предварительная диагностика элементозов по имеющимся у пациента симптомам с учетом провоцирующих факторов. Эта возможность может быть использована как для “грубой” диагностики при отсутствии возможности провести анализ того или иного биосубстрата пациента, так и для лучшей интерпретации результатов анализов. Как известно, высокое количество, например, кальция в волосах при определенных условиях может говорить о его катастрофической нехватке в организме. Поэтому для постановки более точного диагноза необходимо учитывать всю имеющуюся о пациенте информацию.
- Подбор для конкретного пациента списка лекарственных препаратов и биоактивных добавок с учетом степени дисбаланса микроэлементов, имеющихся между микроэлементами синергетических и антагонистических связей, наличия у пациента побочных заболеваний, противопоказаний.
- Составление схемы приема лекарств на неделю, на месяц, на полный курс лечения с учетом сочетаемости отдельных препаратов.
- Подбор рациона (ежедневного, недельного, месячного) с учетом показаний и противопоказаний.
- Формирование отчета для пациента и выдача его на печать.

Таким образом, по выполняемым функциям эта система является справочно-информационно-экспертной системой.

2 Онтологический анализ и синтез. Инструментальные средства

При проектировании современных систем модель предметной области создается на основе онтологического анализа. Как отмечалось в работах [3, 4], можно рассматривать два типа онтологического анализа, в результате которых создаются онтологии предметных областей и методо-ориентированные онтологии. Онтологии первого типа содержат понятия предметной области, и отражают существующие между ними структурные и содержательные связи. Онтологии второго типа представляют имеющиеся в системе средства обработки знаний, логического вывода, выполнения расчетов и т.д. Действие средств, представленных в методо-ориентированных онтологиях над понятиями предметной области, есть процесс онтологического синтеза.

Существует много средств представления онтологий и реализации онтологического синтеза. В работах [4, 5] говорится о применении объектно-ориентированного подхода, о построении онтологий на основе таблиц «объект-свойство» (ТОС) и рекомендуется использовать для представления онтологий неоднородных семантических сетей. Для реализации рассматриваемой системы нами был выбран технологический комплекс Semp-ТАО, который поддерживает объектно-ориентированный стиль представления знаний и предоставляет разработчику удобные универсальные средства конструирования онтологий и осуществления на их основе синтеза: семантические сети, реляционные базы данных, продукционные системы.

3 Онтология нутриентного баланса

Основной задачей на первом этапе разработки любой системы является создание модели предметной области и, прежде всего, ее онтологии. В данной работе в качестве предметной области выбран раздел медицины, который занимается изучением влияния микроэлементов, витаминов на организм человека, а также вопросами нутриологии (науки о здоровом питании). На рисунке 1 представлен фрагмент онтологии нутриентного баланса.

В предметной области выделено шесть основных понятий: *Нутриент*, *Болезнь*, *Фактор*, *Показатель*, *Источник* и *Пациент*. Понятия могут являться базовыми для других понятий. Так, например, *Источник* может быть как *Препаратом*, так и *Продуктом* питания, а *Показатель* порождает такие понятия как *Субъективный* и *Объективный* (показатели), которые, в свою очередь, тоже могут порождать какие-то понятия. Субъективные показатели – это симптомы, которые пациент может наблюдать у себя. Объективные показатели являются результатами анализов или каких-либо обследований.

Над понятиями предметной области определены отношения. Отношение, определенное над базовым понятием, является определенным и для всех порожденных им понятий.

Приведенный фрагмент является ядром онтологии рассматриваемой области. Каждое понятие может иметь свою собственную онтологию, которая будет отображать сложные связи и отношения, соответствующие данному понятию. Рассмотрим более подробно понятие *Болезнь*. Оно порождает три класса: *Дефицит* нутриентов, *Избыток* нутриентов и *Прочие* болезни, которые, не являясь предметом данного исследования, должны, тем не менее, изучаться и учитываться, так как они могут быть как причиной, так и следствием нарушения микроэлементного баланса в организме пациента. Причиной *Болезней* могут быть различные *Факторы* – *Образ жизни* пациента, *Род его занятий* и наличие *Вредных привычек*. *Болезни* могут иметь те или иные *Показатели*. При определенных *Болезнях* могут быть показаны или противопоказаны те или иные *Нутриенты*. И, наконец, *Болезни* имеют непосредственное отношение к *Пациентам* – их носителям.

На основе разработанной онтологии строится модель предметной области, и подбираются средства и схемы онтологического синтеза. Технологическую основу данной онтологии составляет семантическая сеть. Все понятия предметной области – это классы семантической сети. А, собственно, модель предметной области – это совокупность объектов этих классов и экземпляров отношений над ними. Комплекс Semp-ТАО поддерживает объектно-ориентированную парадигму программирования. Классы понятий содержат в себе большую часть существующих между ними связей. Некоторые отношения онтологии представлены отдельными объектами семантической сети – экземплярами отношений (*Быть назначенным (Источник, Пациент)*, *Иметь (Пациент, Показатель)*, *Иметь (Пациент, Болезнь)*). Достаточно большой фрагмент онтологии представляется в виде таблиц. Например, так задается связь между объектами классов *Источник* и *Нутриент*. Таким образом, содержание МЭ и витаминов в тех или иных продуктах или препаратах задается средствами стандартной реляционной базы данных, интеграцию с которой позволяет осуществлять технологический комплекс. Небольшой фрагмент табличного отношения представлен в таблице 1.

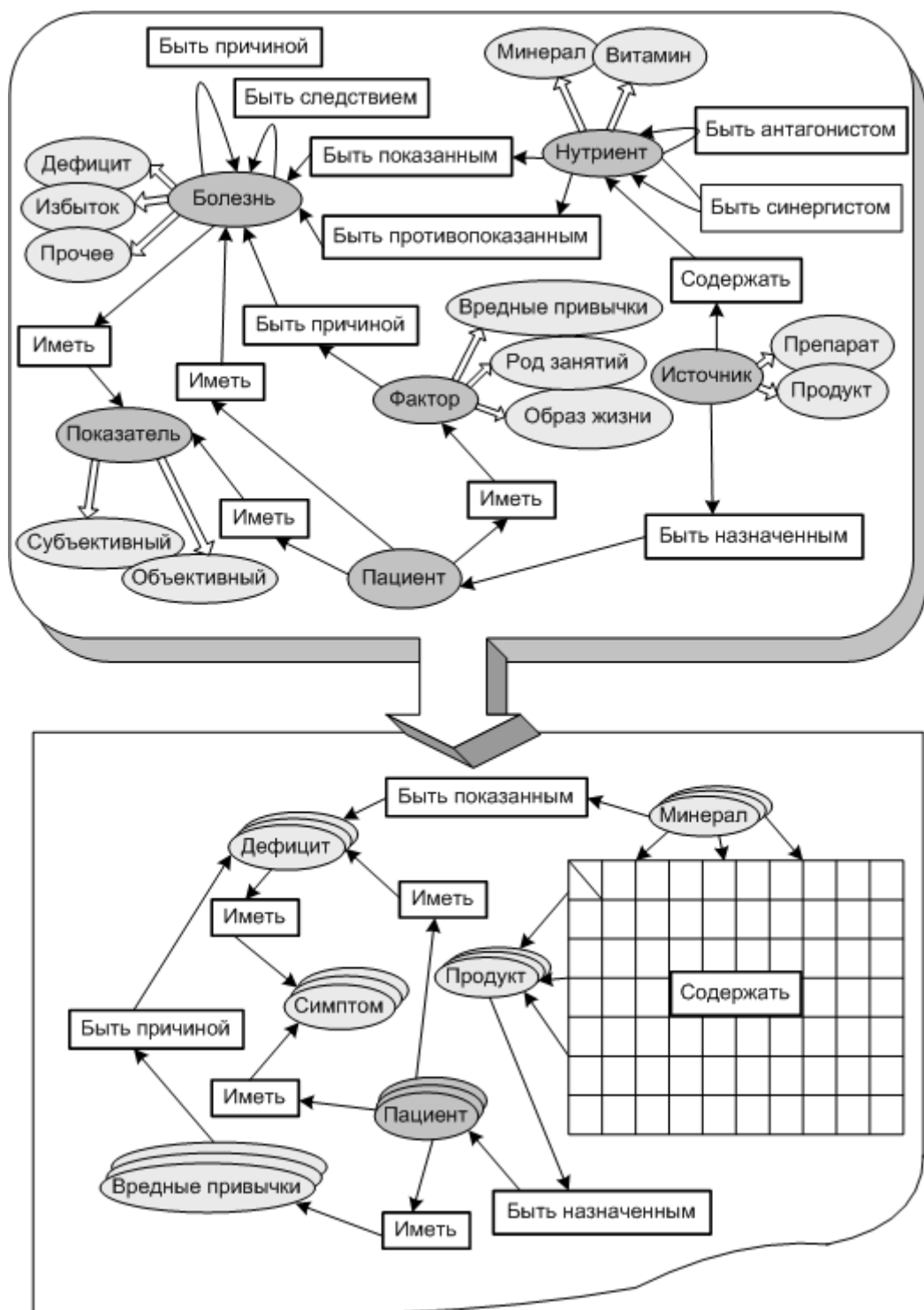


Рисунок 1– Онтология и модель нутриентного баланса организма.

Таблица 1 - Содержание МЭ в продуктах питания.

Продукт/Минерал	Ca	Fe	K	Mg	Na	P	Si	Zn
гречневая крупа	20	6.7	380	200	3	298	81	2.05
икра	75	2	264	141	0	426	0	0
молоко	120	0.07	146	14	50	90	0	0.39
овсяная крупа	64	3.9	362	116	35	349	43	2.68
томаты	14	0.9	290	20	40	26	0	0.2
творог	120	0.3	117	24	44	189	0	0.364
черника	16	7	51	6	6	13	0	0
яблоки	16	2.2	278	9	26	11	0	0.15
яйца	55	2.5	140	12	134	192	0	1.1
...								

Запись: 11 из 11

Процесс онтологического синтеза осуществляется средствами систем продукций, реализованных в Semp-ТАО. Под онтологическим синтезом здесь понимается как процесс интеграции знаний, получение новых знаний из имеющихся, так и получение из представленных в модели предметной области знаний и фактов новых фактов.

4 Пример работы системы

Рассмотрим небольшой пример осуществления онтологического синтеза.

Пусть мы хотим сбалансировать питание для пациента, у которого имеются определенные проблемы со здоровьем.

Сначала в систему вводятся данные о пациенте, и выбирается режим работы системы. Получение списка нужных продуктов будет происходить по следующей схеме (см. рисунок 1), которая содержательно сильно упрощена и служит лишь для демонстрации средств онтологического синтеза:

- Если в семантической сети будут обнаружены факты наличия у пациента показателей и факторов определенного элементоза (Дефицита или Избытка какого-либо Минерала), то факт наличия заболевания будет также занесен в сеть.

*Иметь (Пациент, Вредная_Привычка), Иметь (Пациент, Симптом),
 Быть_причиной (Вредная_привычка, Дефицит),
 Иметь (Дефицит, Симптом) ⇒
 Иметь (Пациент, Дефицит);*

- Далее, при наличии факта Дефицита определенного Минерала из таблицы будут выбраны продукты, в которых он содержится.

*Иметь (Пациент, Дефицит), Быть_показанным (Минерал, Дефицит),
 Содержать (Продукт, Минерал) ⇒
 Быть_назначенным (Продукт, Пациент);*

Составление схем приема препаратов, дневного (недельного) рациона и меню – задачи достаточно сложные, творческие и трудоемкие. Их обсуждение выходит за рамки данной работы.

Заключение

В статье рассмотрена задача разработки системы, предназначенной для лечащего врача и ориентированной на информационную поддержку и сопровождение курса лечения заболеваний, связанных с нарушением микроэлементного и витаминного баланса организма. Показана важность и актуальность этой задачи.

На основе онтологического анализа создана модель предметной области. Выбраны средства реализации системы и показана их адекватность поставленной задаче.

Выполнена реализация некоторых компонентов онтологии и модели предметной области, опробованы схемы онтологического синтеза, реализующего функциональные возможности системы.

Список литературы

- [1] Загорюлько Г.Б., Гончар А.М. Подход к разработке экспертной системы корректировки микроэлементного состава организма на основе данных рентгенофлуоресцентного анализа волос // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, 17-21 июня 2003 г.). - Самара: СНЦ РАН, 2003. С. 242-247.
- [2] Закаса Н.П., Колмогоров Ю.П., Гончар А.М., Шелпакова И.Р. Возможности РФА СИ и АЭС анализа при определении микроэлементов волос // XIV Российская конференция по использованию синхротронного излучения (Новосибирск, июнь 2002 г.).
- [3] Виттих В.А. Онтологический анализ и синтез при управлении сложными открытыми системами // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, 17-21 июня 2003 г.). - Самара: СНЦ РАН, 2003. С. 50-60.
- [4] Кузнецов С.В., Смирнов С.В. Методо-ориентированные онтологии и описание объектной динамики. // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, 17-21 июня 2003 г.). - Самара: СНЦ РАН, 2003. С. 108-114.
- [5] Смирнов С.В. Онтологический анализ в системах компьютерного моделирования // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, 17-21 июня 2003 г.). - Самара: СНЦ РАН, 2003. С. 102-107.
- [6] Загорюлько Ю.А., Попов И.Г. Описание сложных предметных областей на основе интеграции средств представления знаний // Труды международного семинара Диалог'97 по компьютерной лингвистике и ее приложениям. - Москва, 1997. С. 110-115.

ИСПОЛЬЗОВАНИЕ НЕДООПРЕДЕЛЕННЫХ ВЫЧИСЛЕНИЙ ДЛЯ РЕШЕНИЯ ОПТИМИЗАЦИОННЫХ И ОБРАТНЫХ ЗАДАЧ НА ГРАФАХ

В.В. Тарасевич

Российский НИИ институт искусственного интеллекта
630090, Новосибирск, пр. Лаврентьева, 6, Россия
tvv@iis.nsk.su
тел.: +7 (3832) 34-29-91, факс +7 (3832) 32-83-59

Ключевые слова: *недоопределенность, вычисления, управление по данным, вычислительная модель, задача на графах, система, трубопровод, параметр*

Abstract

A new approach is suggested for the solution of direct and inverse problems in hydraulic engineering with imperfect data under presence of set of constraints of various kinds. This approach is based on artificial intelligence techniques.

The paper summarizes many years' results on adaptation of artificial intelligence technique for engineering problem solution. The applications of such techniques as the subdefinite computational models, the constraint programming, etc, are considered in the paper. Their advantage consists in the following: they a) allow use the subdefinite and incomplete data; b) allow to integrate the heterogeneous constraints by a natural way; c) allow equally simply to consider both direct and inverse problems; d) require description of computational model only, what eliminates the need for algorithm design.

The using of aforesaid AI-techniques is illustrated by examples of solution the design problems of pipeline and piping network. The problem of piping networks' optimal designing is among them. These problems belong to category of graph-defined problems. The classification of the problems is suggested, which considers the direct well-conditioned and ill-conditioned problems and inverse problems. The results of numerous computational experiments are discussed.

Введение

Многие технические задачи, особенно связанные с природной средой, оперируют не с точно определенными исходными данными, а некоторым интервалом вероятных значений. Строго говоря, результаты расчетов также являются некоторыми интервалами. Обычно пользователь при расчетах использует некоторое среднее значение (представителя из интервала) и результат получает тоже в виде некоторого единичного значения (представителя из интервала результатов), что неявно подразумевает оценку погрешности на выходе, которая может быть существенной.

Характерной особенностью таких задач является наличие множества различного вида ограничений: технологических, экологических, экономических и др. Это обстоятельство значительно осложняет процесс решения задачи, повышает степень ее сложности, делает алгоритм решения нерегулярным. Другая проблема состоит в том, как интегрировать в процесс решения опыт пользователя и использовать имеющиеся у него знания индивидуальных особенностей задачи, дополнительные данные и т.п.

В последнее время на практике стали играть все более значительную роль различного вида обратные и плохо обусловленные задачи, возникающие при поиске оптимального варианта, решении задач оптимального управления, идентификации системы, задач прогноза и т.д. Обратные и плохо обусловленные задачи имеют на порядок большую сложность вычислений по сравнению с прямыми задачами, и требуют специальной техники нахождения решения. Не-

смотря на многочисленные усилия исследователей, в данном направлении имеется еще много нерешенных проблем.

Использование технологии недоопределенных вычислений может преодолеть многие из вышеотмеченных трудностей. Легкость учета различных ограничений, декларативность входного языка, отсутствие различий между «входными» и «выходными» параметрами, гибкость управляемого по данным процесса вычислений делает их весьма привлекательными для решения как прямых, так и обратных задач [1, 2].

В настоящей работе применение технологии недоопределенных вычислений [3] продемонстрировано на примере задач о расчете трубопровода и трубопроводной сети, которые относятся к категории задач, определенных на графе.

1. Математическая постановка задачи

1.1. Структура трубопроводной системы

Под трубопроводной системой будем понимать множество соединенных некоторым образом трубопроводов; в точках соединения труб, называемых узлами, могут находиться различные устройства (клапаны, насосы, регуляторы, фильтры и другая гидроарматура). Структуру системы принято изображать в виде некоторого ориентированного графа Γ , у которого вершины соответствуют узлам сети, а ребра – трубам [4]. Примерами таких систем могут служить сети теплоснабжения, водоснабжения, системы подачи топлива, технологические линии, системы магистральных нефте- и газопроводов, и т.п.

Пусть трубы системы пронумерованы от 1 до S , а узлы – от 1 до m , нижний индекс i соответствует трубопроводу, а верхний индекс j соответствует узлу системы. Будем обозначать множество номеров всех труб, примыкающих к узлу j , через R^j ; а η_i^j будет обозначать параметр инцидентности: $\eta_i^j = -1$, если i -я труба выходит из узла j ; и $\eta_i^j = 1$, если i -я труба входит в узел j .

1.2. Уравнения для трубопровода

Основным уравнением, которое описывает связь между расходом Q в трубе и разностью энергетических потенциалов на концах трубы (в качестве такового выступает напор H) будет известное соотношение Дарси-Вейсбаха [4, 5]:

$$(1) \quad H_i(0) - H_i(L) = A_i L_i |Q_i| Q_i.$$

Здесь $H(0)$ – напор в начале трубы; $H(L)$ – напор в конце трубы; L – длина трубы, A – удельное сопротивление, которое определяется по формуле (индекс i здесь и далее для простоты опущен)

$$(2) \quad A = \frac{8\lambda}{g\pi^2 d^5},$$

где d – диаметр, λ – коэффициент гидравлического трения:

$$(3) \quad \lambda = \lambda(k/d, \mathbf{Re}).$$

Здесь k – шероховатость стенок трубы, $Re = 4Q/\pi d\nu$ – число Рейнольдса, где ν – кинематическая вязкость жидкости [5]. Заметим, что шероховатость k является не точно заданной, а интервальной величиной. Для определения λ существует большое множество формул (см., [5, 6]). В качестве наиболее популярных можно привести явную формулу Альтшуля [5]:

$$(4) \quad \lambda = 0.11 \left(\frac{k}{d} + \frac{68}{\mathbf{Re}} \right)^{0.25},$$

и неявную формулу Колбрука - Уайта [6]:

$$(5) \quad \frac{1}{\sqrt{\lambda}} = -2 \log \left(\frac{k}{3.7d} + \frac{2.51}{\text{Re} \sqrt{\lambda}} \right).$$

1.3. Граничные условия

Пусть \bar{Q}^j и \bar{H}^j обозначают вектора, составленные соответственно из значений расходов и напоров на концах труб, примыкающих к узлу j :

$$(6) \quad \bar{Q}^j = \{Q_i\}, \quad \bar{H}^j = \{H_i^j\}, \quad i \in R^j,$$

где $H_i^j = H_i(0)$, если труба i выходит из узла j ; и $H_i^j = H_i(L_i)$, если труба i входит в узел j .

В общем виде уравнения, описывающие функционирование узлов, можно описать следующим образом:

$$(7) \quad \bar{F}^j(\bar{Q}^j, \bar{H}^j, \bar{\Lambda}^j) = 0,$$

где $\bar{\Lambda}^j$ – вектор собственных параметров узла (также неизвестная величина).

Например, для соединения труб имеем:

$$(8) \quad H_i^j = h^j \quad \forall i \in R^j, \quad \sum_{i \in R^j} \eta_i^j Q_i^j = q^j,$$

где h^j – давление в узле j ; q^j – утечка (отбор) из узла j . В этом случае вектор собственных параметров $\bar{\Lambda} = (h^j, q^j)$. Частным случаем (8) для висячих вершин (т.е. когда $R^j = \{i\}$) будет задание напора h^j или расхода q^j в узле.

Важным частным случаем граничных условий является ситуация, когда в висячих вершинах заданы напоры или расходы, а во внутренних вершинах заданы условия сопряжения в виде балансовых соотношений (8).

2. Классификация задач

Так как параметры гидравлического режима (расходы и напоры) обусловлены параметрами трубопроводной системы (диаметры, шероховатость и т.д.), естественно считать параметры трубопроводной системы входными параметрами задачи, а параметры гидравлического режима – выходными.

Задачи, для которых по заданным входным параметрам надо определить выходные параметры будут называться прямыми.

Когда же, наоборот, по заданным выходным параметрам надо определить входные, такие задачи будем классифицировать как обратные задачи.

Прямая задача будет называться хорошо обусловленной, если система уравнений (1) – (3), (7) имеет единственное решение при выполнении условия

$$(9) \quad A_i \equiv 0 \quad \text{для всех } i = 1, \dots, S.$$

В противном случае задачу назовем плохо обусловленной. Хорошая обусловленность означает фактически требование слабой чувствительности задачи к входным параметрам.

Допустим, требуется определить гидравлический режим в трубопроводной системы, когда узлы описываются граничными условия (8), при этом в r первых вершинах заданы напоры h^j , где $j = 1, \dots, r$; а во всех остальных вершинах заданы отборы q^j , где $j = r+1, \dots, m$. Следуя терминологии [7], такую задачу будем называть канонической.

Для $r = 1$ каноническая задача будет типичным примером прямой хорошо обусловленной задачи, при этом условие (9) делает систему уравнений (1) – (3), (7) линейной, а требование хорошей обусловленности означает невырожденность ее матрицы. При $r > 1$ каноническая задача будет плохо обусловленной, т.к. имеет только вырожденное решение при условии (9).

3. Оптимизационные задачи

3.1. Задача о подборе диаметров

В случае одиночного трубопровода ($S=1$) задача сводится к отысканию удельного сопротивления A по (1), зная $H(0)$, $H(L)$, и Q . Далее по формулам (2) – (3) находится искомое значение d . Но решение затрудняется тем, что на самом деле значения диаметров принимают дискретные значения согласно промышленному сортаменту: $d \in D$, где D – множество значений диаметров, $D = \{D_1, D_2, \dots, D_m\}$, причем $D_1 < D_2 < \dots < D_m$.

На практике обычно найденное из решения (1) – (3) значение d округляют в большую сторону до ближайшего номенклатурного значения из D . Но полученное таким образом решение, строго говоря, уже не будет удовлетворять равенству (1). Для корректности постановки задачи вместо соотношения (1) используется энергетическое неравенство, ограничивающее пропускную способность трубопровода снизу:

$$(10) (H_i(0) - H_i(L_i)) \cdot Q_i \geq A_i L_i |Q_i|^3$$

В случае системы труб для решения задачи уже недостаточно чисто гидравлических соотношений (1) – (10). Для замыкания задачи применяется требование минимальной стоимости (в том или ином смысле) трубопроводной системы, например:

$$(11) \sum_i S_i(d_i) L_i \rightarrow \min$$

где $S_i = S_i(d_i)$ – стоимость одного погонного метра i -ой трубы. В частности, (11) превращается в требование минимальной материалоемкости при $S_i(d_i) = \pi \delta_i (\delta_i + d_i)$, где δ_i – толщина стенок трубы.

Задача формулируется так: найти дискретные значения d из множества D , удовлетворяющие уравнениям (2) – (3), (7) и неравенству (10), для которых функционал (11) достигает минимума, т.е. задача сводится к оптимизационной задаче.

3.2. Определение шероховатости k (задача идентификации, или калибровки)

Для одиночного трубопровода ($S=1$), аналогично предыдущему случаю, зная A , легко определить k , решая совместно (2) и (3). В случае системы труб для замыкания задачи применяется требование минимального отклонения наблюдаемых величин от расчетных:

$$(12) \sum_i \alpha_i (Q_i - \bar{Q}_i(k_i))^2 + \sum_j \sum_{i \in R^j} \beta_i^j (H_i^j - \bar{H}_i^j(k_i))^2 \rightarrow \min$$

где величины с чертой сверху – расчетные значения при заданной шероховатости, а без черты – наблюдаемые значения, α и β – весовые коэффициенты.

Требуется найти положительные k , удовлетворяющие соотношениям (1) – (3), (7), (12).

4. Методика расчета

Подход базируется на представлении решаемой задачи в виде некоторой семантической сети [2], при этом отдельным сущностям (трубопроводам и узлам системы) будут соответствовать объекты такой сети, а связи между объектами будут описывать зависимости между этими сущностями в соответствии с (6). Соотношения (1) – (3) будут служить внутренней вычислительной моделью объекта типа «труба», а соотношения (7) – внутренней вычислительной моделью объекта типа «узел». Левые части (11) или (12) будут представлять собой глобальную вычислительную модель (в случае обратных задач, см. п.3). На начальной стадии часть параметров (исходные данные задачи) задаются точными значениями (или уточненными интервалами), а остальные параметры (искомые величины) являются недоопределенными величинами, которые подлежат уточнению в процессе активации семантической сети [2].

Активация семантической сети происходит по потоковому принципу: уточнение значений параметров одних узлов продуцирует изменение параметров других, связанных с ними узлов, и т.д., вплоть до полного успокоения сети. Полученные уточненные значения (вообще говоря, интервальные) тех параметров, которые относятся к искомым величинам, и есть «решение» задачи (в классическом смысле).

Если полученные уточненные интервалы искоемых величин удовлетворяют требованиям пользователя, то задача считается «решенной». Если же некоторые интервалы слишком велики, то это подсказывает пользователю, где и какая дополнительная информация требуется для продолжения процесса уточнения недоопределенных параметров (решения задачи).

Применение технологии недоопределенных вычислений к решению прямых задач (см. п. 2) описано и проиллюстрировано примерами в вышеупомянутых работах [1, 2]. Заметим, что в рамках парадигмы недоопределенных вычислений нет разницы между использованием формул (4) и (5), в то время как для традиционных вычислений неявность формулы (5) служит препятствием ее применения.

5. Примеры расчетов трубопроводных систем

Рассмотрим пример напорной оросительной системы, ориентированный граф которой изображен на рисунке 1.

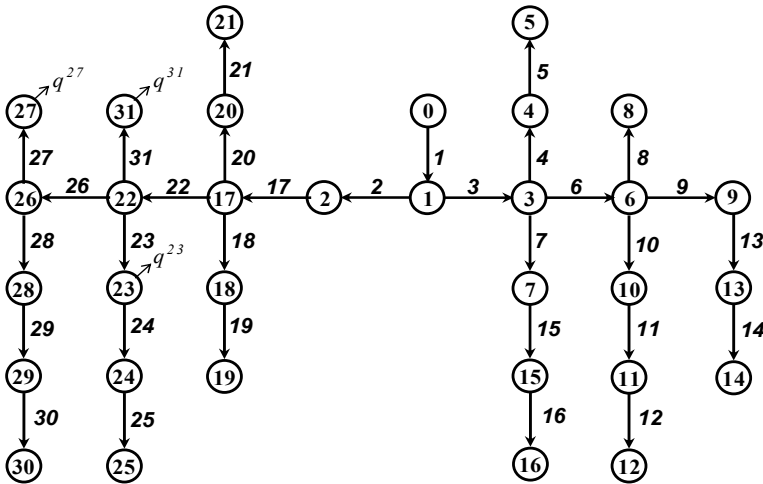


Рисунок 1 – Граф напорной оросительной сети.

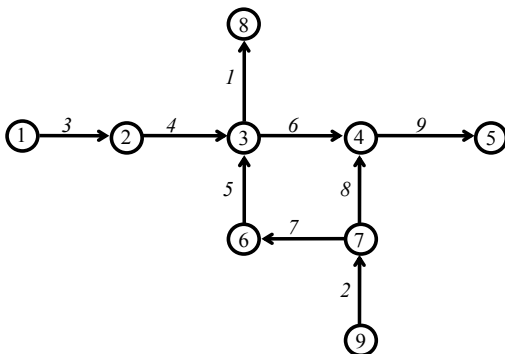


Рисунок 2 – Схема сети водоснабжения

Здесь номера в кружках – номера узлов, а номера возле дуг – номера труб. В узле 0 расположена насосная станция. Все другие узлы представляют собой соединения труб, описываемые уравнениями (8). Расчет параметров потоков в системе при заданном головном напоре h^0 и отборах q^j в некоторых узлах является хорошо обусловленной задачей.

Обратная задача по подбору диаметров была решена при условии минимальной материалоемкости, когда заданы головное давление, рабочие напоры и требуемые расходы на выходе при соблюдении ряда технологических ограничений.

Схема системы водоснабжения заводского цеха представлена на рисунке 2. В узлах 1, 8, 9 заданы давления и в узле 5 задан расход. Внутренние узлы 2, 3, 4, 6, 7 представляют собой соединения труб без утечек, описываемые уравнениями (8).

Неизвестными параметрами потока являются расходы Q_i , и давления $H_{0,i}$ и $H_{L,i}$, которые удовлетворяют уравнениям (1) – (3) для каждой i -той трубы. Сопротивления труб A_i заданы. Задача расчета гидравлических параметров потока в такой системе будет плохо обусловленной; наличие замкнутого контура (узлы 3, 4, 7, 6) значительно усложняет процесс решения.

Заключение

Задачи расчета трубопроводных систем классифицируются на прямые хорошо и плохо обусловленные задачи и обратные задачи. Установлено, что применение недоопределенных вычислений весьма эффективно для хорошо обусловленных задач, а также дает удовлетворительные результаты также для систем без колец. Более того, применение этих вычислительных технологий в данных случаях обеспечивает новые возможности и расширяет класс решаемых задач, позволяя работать с неточно и неполно заданными исходными данными, получая на выходе решения в виде интервалов допустимых значений неизвестных. Для плохо обусловленных задач на графах с кольцами представляется более эффективным путь использования недоопределенных вычислительных моделей в комбинации с другими технологиями вычислений, взаимно усиливая их сильные стороны. Применение технологии недоопределенных вычислений для обратных задач, таких как задача о подборе диаметров или задачи идентификации представляется весьма перспективным.

Данный подход легко обобщается на решение других задач, определенных на графах, например, для расчета распределения токов и напряжений в электрических сетях, для расчета транспортных потоков и др.

Список литературы

- [1] Narinjani, A.S., Popov, I.G., Poltoratzky, A.V., Tarasevich, V.V. and Zagorulko, Y.A. Application of subdefinite data driven computation technique to the water network calculation, Water Network Modelling for Optimal Design and Management, Proc. CWS 2000, Sept 2000, eds. by D. Savic and G. Walters, Centre for water systems, University of Exeter, UK, pp. 251-261.
- [2] Загорюлько Ю.А., Попов И.Г., Тарасевич В.В. Моделирование потоков в сложных системах с применением интеллектуальных технологий. //Проблемы управления и моделирования в сложных системах: Труды I Международной конференции – Самара: Самарский научный центр РАН, 1999. – с.39-44.
- [3] Нариньяни А.С. Недоопределенность в системах представления и обработки знаний // Изв. АН СССР. Техн. кибернетика. 1986, №5. С. 3-28.
- [4] Меренков А.П., Хасилев В.Я. Теория гидравлических цепей. - М.: Наука, 1985.
- [5] Альтшуль Д.А. Гидравлические сопротивления. – М.: Наука, 1982.
- [6] Souza, J., da Conceicao Cunha, M. and Sa Marques, A. An explicit solution to the Colebrook-White equation through simulated annealing. Water Industry Systems: modelling and optimization applications, eds. by D. Savic and G. Walters, vol. 2, Research Studies Press Ltd., UK, 1999, pp. 347-355.
- [7] Сухарев М.Г., Карасевич А.М. Технологический расчет и обеспечение надежности газо- и нефтепроводов. – М.: ГУП Издательство «Нефть и газ» РГУ нефти и газа им. И.М. Губкина, 2000.

ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ДЛЯ РАЗРАБОТКИ МУЛЬТИАГЕНТНЫХ СИСТЕМ ПРОМЫШЛЕННОГО МАСШТАБА

В.В. Андреев¹, С.В. Батищев², К.В. Ивкшкин², Т.В. Искварина², П.О. Скобелев²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²ООО НПК «Маджента Девелопмент»

443110, Самара, ул. Осипенко, 1а
andreev@magenta-technology.ru

тел: +7 (8462) 70-66-84; 70-66-85

Ключевые слова: онтология, агент, предприятие, Java, J2EE, мультиагентные системы, виртуальный рынок

Abstract

The article describes tools for development enterprise ready Multiagent systems, which are based on J2EE technology. Article describes notions and approaches of ontologies managing, virtual market and requirements for enterprise ready Multiagent system.

Введение

Среди существующих подходов к построению мультиагентных систем (МАС) можно выделить два класса систем:

- Распределённые МАС с “высокоинтеллектуальными” агентами.
- МАС основанные на “групповом разуме” (swarm intelligence).

Первый класс систем рассчитан на распределённые системы, в которых на отдельно взятом сервере существует сравнительное небольшое количество агентов (единицы, десятки) и каждый из агентов является сложным программным объектом со сложными алгоритмами и возможностью межмашинного обмена данными и командами.

Втором класс систем рассчитан на большое количество агентов работающих на одном сервере (тысячи, десятки тысяч). Каждый агент в системе не является сложным объектом и, как правило, реализует простые алгоритмы переговоров, однако в результате взаимодействия агентов возможно решение сложных задач, которые каждый агент по отдельности решить не в состоянии. Коммуникация между агентами на разных серверах не используется широко в таких системах в виду большой интенсивности переговоров между агентами и как следствие – уменьшения быстродействия систем из-за потерь на межмашинном обмене данными.

Оба класса систем могут быть применены к решению задач оперативной обработки информации в процессах принятия решений, где необходимо обеспечить процедуры принятия и пересмотра решений в условиях высокой динамики среды, учёта предпочтений каждого участника взаимодействия и использования формализованных знаний ЛПР при принятии решений. Будем называть такие системы открытыми мультиагентными системами оперативной обработки информации для поддержки процессов принятия решений (ОМАС ППР) [1].

Для эффективного создания ОМАС ППР масштаба предприятия необходимо использование инструментальных средств, которые упрощают и ускоряют процесс разработки, тестирования и внедрения систем. Такие средства должны обеспечивать построение систем, обладающих следующим свойствами:

- Масштабирование: возможность модульного наращивания системы в рамках унифицированной архитектуры;
- Надежность: способность восстановления после сбоев и гарантированная сохранность информации;
- Безопасность: обеспечения контроля доступа к данным и исключение несанкционированного доступа к данным;
- Кроссплатформенность: работа клиент-серверной архитектуры на различных операционных системах;
- Интеграция и совместимость: взаимодействие с другими программными решениями и сервисами;
- Распределенность: возможность работы в многомашинном варианте, обмен данными и командами в локальных и глобальных информационных сетях.

С точки зрения разработки прикладных МАС, инструментальные средства должны обеспечивать следующие требования [2]:

- Независимость агентов, обмен информацией между агентами, использование формализованных знаний в процессах принятия решений
- Эффективное расширение функциональных и сервисных возможностей предлагаемых решений
- Разработку решений в кратчайшие сроки и уменьшение материальных затрат реализации систем

Инструментальные средства, рассмотренные в статье [2] удовлетворяют основным требованиям к ОМАС ППР. Однако данная система была не полностью подготовлена для использования в масштабе предприятия, т.к. была ориентирована на работу на изолированной ЭВМ, поддерживала онтологии сравнительно небольших размеров и т.д.

1 Предлагаемый подход и архитектура системы

Развитием системы описанной в статье [2] стала новая линейка инструментальных средств компании Маджента, основывающихся на платформе Java 2 Platform, Enterprise Edition (J2EE). Платформа J2EE предназначена для разработки приложений и программных решений масштаба предприятия. Реализации спецификаций J2EE предоставляют поддержку и модель для построения систем отвечающих требования безопасности, кроссплатформенности и совместимости. Используя платформу J2EE, компания Маджента развивает собственные инструментальные средства для разработки МАС масштаба предприятия.

На рисунке 1 изображена физическая архитектура систем, построенных на базе инструментальных средств компании Маджента.

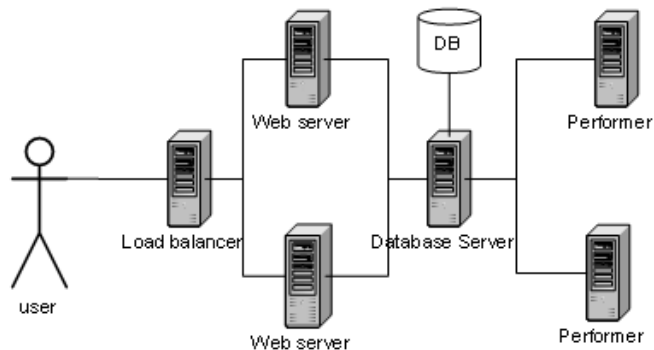


Рисунок 1 – Физическая архитектура систем

Выделяется уровень представления (web серверы), уровень данных (база данных, сервер баз данных) и уровень бизнес-логики (сервера исполнители). Такая структура позволяет увеличивать количество серверов обрабатывающих запросы без изменения архитектуры, а так же обеспечивается параллельное вычисление задачи на различных серверах.

Рассмотрим логическую архитектуру систем построенных на базе инструментальных средств:

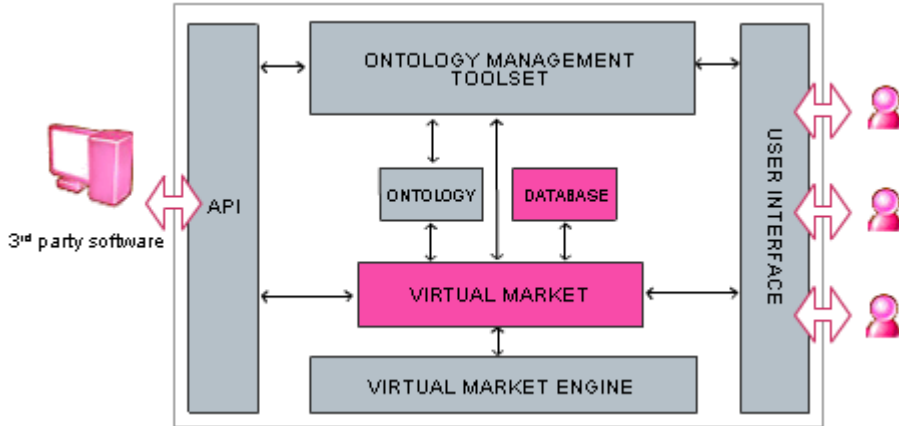


Рисунок 2 – Логическая архитектура систем

Основными составляющими частями инструментария являются Ontology Management Toolset (OMT) и Virtual Market Engine (VME).

OMT является инструментальным средством построения и использования онтологий. OMT позволяет:

- проектировать онтологии в виде семантических сетей. Используя OMT, разработчик и пользователь может создавать и редактировать онтологии, специфицируя свои концепты (классы) и устанавливая связи между ними, а также формируя сценарии действий;
- создавать сцены и работать с ними. Сцена – это набор экземпляров классов, описывающих некоторую ситуацию в мире. Каждый из экземпляров набора связан с некоторым концептом онтологии отношением вид-род;
- обеспечить описание классов агентов и правил принятия решений в онтологии;
- предоставление функций по работе с экземплярами классов и экземплярами отношений;
- осуществлять навигацию по семантическим сетям онтологий и сцен (получение связанных сущностей, проверка наличия пути и т.д.). При этом функции по навигации доступны как разработчикам с помощью программного интерфейса приложения (API), так и в скрытых агентов.
- хранить онтологии и сцены в реляционной базе данных;
- импорт/экспорт онтологий в XML.

Рассмотрим понятие виртуального рынка (Virtual Market). Виртуальный рынок – это среда взаимодействия агентов, которая включает следующие части:

- Спецификация (описание) типов агентов и структуры агентов
- Спецификация взаимодействия агентов:
 - Спецификация типов сообщений и структуры сообщений
 - Протоколы взаимодействия агентов
- Программная реализация виртуального рынка

- Дополнительные программные компоненты, специфичные для конкретного виртуального мира (например, объект расписания, предоставляющий функции по низкоуровневой работе с расписанием ресурсов)

Виртуальный рынок проектируется и реализуется для решения определённого рода задач и является неотъемлемой частью МАС. Приведем пример виртуального рынка - виртуальный рынок транспортной логистики, использующий такие понятия как заказ, ресурс и предназначенный для решения задач планирования и перепланирования движущихся ресурсов.

2 Ontology Management Toolset

При построении программных систем, все знания, необходимые системе для функционирования, можно разделить на две части:

- знания предметной области;
- знания, относящиеся к методу решения задачи или принятия решений.

Знаниями предметной области является, например, то, что в мире бывают заказы и ресурсы, что ресурсами бывают грузовики, корабли, грузы, маршруты и прочие знания, не зависящие от конкретного приложения, типов переговоров агентов и методов решения.

Знаниями, относящимися к методу решения задачи, является все, что связано с принятием решений и переговорами агентов.

Для описания знаний предметной области используется онтология предметной области (Domain ontology). Соответственно, знания, относящиеся к методу решения задачи, описываются в онтологии виртуального рынка (Virtual Market Ontology). Выделение этих уровней позволяет отделить знания предметной области от знаний об организации конкретного виртуального рынка. Благодаря такому разделению можно создать онтологию некоторой предметной области и использовать ее независимо в нескольких решениях.

Рассмотрим двухуровневую схему построения МАС: первый уровень – универсальная МАС, рассчитанная на решение некоего класса задач (например, транспортная логистика), на втором уровне на базе МАС первого уровня строится система для решения некоторой конкретной задачи в рамках заявленного класса задач.

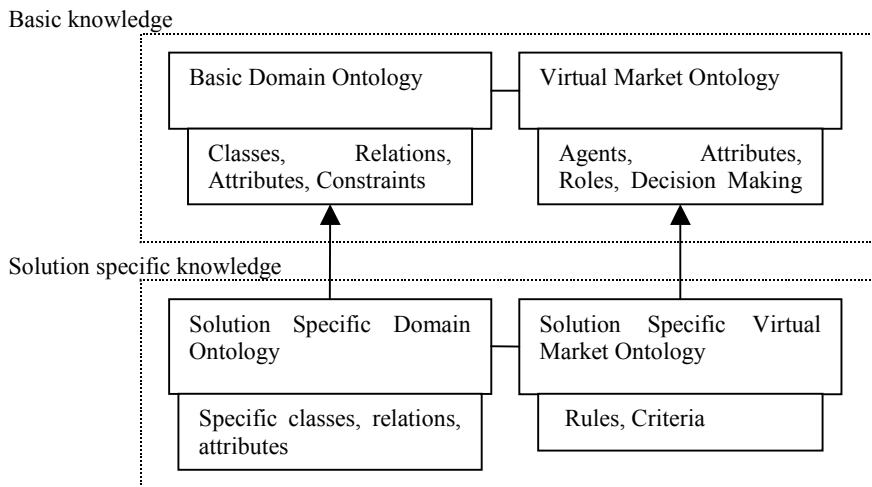


Рисунок 3 – Типы онтологий

В случае такой схемы построения МАС предлагается следующий принцип разделения формализованных знаний (см. рисунок 3). Выделяется базовый слой онтологий, который опи-

сывает знания необходимые для функционирования системы на общем, универсальном уровне. Так, для транспортной логистики, базовая онтология предметной области содержит понятия ресурс, груз, маршрут. Базовая онтология виртуального мира содержит классы агентов, такие как агента заказа, агент груза и описание ролей агента. Для конкретного решения на базе такой МАС разработчики и пользователи могут создать свои онтологии предметной области и виртуального мира, расширив тем самым знания базовых онтологий. В расширенной онтологии предметной области для задач морской логистики можно описывать суда, для систем грузоперевозок по суше – грузовики, трейлеры, тягачи и т.д. В расширенной онтологии виртуального рынка декларируются правила принятия решений в виде скриптов и описываются критерии агентов.

Такая схема позволяет создавать несколько наборов расширенных онтологий для одного набора базовых онтологий, т.е. позволяет добавлять функциональность в решения для различных клиентов на основе одного универсального продукта путем создания новых онтологий. Такой подход позволяет экономить материальные средства и время на разработку конечных продуктов.

При создании онтологий возможно использование следующих понятий и концептов:

- Класс объекта – понятие, описывающее некоторую совокупность объектов одного рода.
- Класс отношения – понятие, определяющее семантическую связь между классами в онтологии.
- Атрибут класса объекта или отношения - является качественной или количественной характеристикой понятия, отражаемого концептом.
- Класс агента – понятие, описывающее тип агентов виртуального рынка, а именно набор ролей, в которых этот агент может выступать в переговорах.
- ОМТ позволяет пользователям задавать правила действий агентов и сценарии в виде скриптов.

2.1 Особенности реализации

Для поддержки исполнения скриптов и увеличения быстродействия работы со сценой используется компиляция онтологий в java-классы, которые описывают экземпляры концептов. ОМТ реализует два основных способа хранения онтологий – в виде XML и с использованием реляционной базы данных. XML является удобным форматом хранения и обмена онтологиями, база данных позволяет хранить и оперировать большими объемами данных, что характерно для ряда предметных областей и применений [4]. Также ОМТ поддерживает систему безопасности и разделения прав пользователя. Данная система реализована на основе Java Authentication and Authorization Service (JAAS) и поддерживает группы пользователей, назначение прав редактирования и просмотра онтологий.

3 Virtual Market Engine

Основное предназначение VME заключается в предоставлении разработчикам программного каркаса, структуры и набора библиотек для эффективного построения виртуальных рынков различной сложности в прикладных МАС.

3.1 Подсистема обмена сообщениями

Система обеспечивает синхронную или асинхронную, а также контекстную или бесконтекстную передачу сообщений между агентами. Простым методом обмена сообщениями является асинхронный, бесконтекстный способ. Этот способ позволяет вести переговоры в условиях неопределённого времени ответа и прихода сообщения, однако требует от разработчика дополнительных усилий по разработке в случае сложных, параллельно идущих переговоров с несколькими участниками. Контекстный метод обеспечивает автоматическую диспетчеризацию сообщений в рамках нескольких протоколов, что обеспечивает возможность удобной ор-

ганизации параллельных переговоров по одинаковым или различным протоколам. Также подсистема обмена сообщениями также позволяет рассылать сообщения группам агентов. В разработке находится поддержка межмашинного варианта обмена сообщениями.

3.2 Подсистема контекстов

Агентный контекст определяет состояние (значения параметров агентов, установленные связи между агентами) сообщества агентов, описывающее некоторую ситуацию.

Основное свойство контекстов заключается в возможности откатов и накатов, т.е. возможности сохранения изменений состояния с последующим откатом до первоначального состояния или наката (операция обратная накату – применение сохраненных изменений). Подсистема контекстов позволяет хранить иерархическую структуру (а в более сложных случаях и сетевую) изменений, так называемое дерево изменений, где предком контекста является контекст (ситуация) который предшествовал данному контексту. Таким образом, для получения некой ситуации необходимо последовательно применить изменения всех предков контекста и затем – изменения самого контекста.

Так как каждый контекст точно описывает ситуацию в агентом сообществе, то такой подход позволяет реализовать вариант многомашинной, параллельной обработки задач и нахождения решения в системах поддержки принятия решений. Необходимо отметить возможность сохранения и восстановления объектов изменений в формат XML, что упрощает поддержку параллельных вычислений в многомашинном варианте, а также масштабирование в случае больших агентных сообществ.

3.3 Диспетчер исполнения

Диспетчер исполнения обеспечивает квазипараллельное исполнение агентов, что является важной особенностью МАС. В основу системы положен принцип невытесняющей многозадачности, при котором агент является активным до тех пор, пока он сам, по собственной инициативе, не отдаст управление диспетчеру исполнения для того, чтобы тот выбрал из очереди другого, готового к активизации агента.

Каждый агент может содержать несколько активных сущностей – ролей (или процессы [1]). Каждая роль может быть построена по принципу конечной машины состояний, что обеспечивает возможность диспетчеризации исполнения агентов и ролей. Весь жизненный цикл роли разбивается на отдельные процедуры (подпрограммы) или состояния. Так, например, после рассылки сообщений адресатам, агент возвращает управление диспетчеру и переходит в состояние ожидания ответов. Диспетчер будет активизировать заданное состояние у агента по наступлению некоторого условия (в данном случае – приход сообщения).

Такая система диспетчеризации рассчитана на режим работы большого количества агентов в серверном варианте, в Enterprise Java Beans (EJB) окружении, где создание и управление потоками управления (нити, threads) запрещено.

4 Принципы конструкции агента

В конструкции агента можно выделить три основных уровня (рисунок 4):

- Уровень коммуникации. На этом уровне выделяются компоненты и функциональные возможности поддержки коммуникации между агентами. Классическим способом взаимодействия агентов – обмен сообщениями. VME обеспечивает асинхронный/синхронный обмен сообщениями между агентами и обмен сообщениями по протоколу (то есть поддержку контекста переговоров).
- Уровень данных. На этом уровне изолируются все свойства и данные агентов: память, опыт, значения онтологических атрибутов и т.д. Уровень данных поддерживает управление историей изменений: отмена изменений, сохранение изменений как отдельной сущности, применение изменений.

- Уровень поведения (управления): уровень поведения является активной составляющей агента. На этом уровне агента определяются роли агента. Роль агента – независимая активная сущность, которая может:
 - Инициировать отсылку сообщения
 - Ждать (находится в состоянии ожидания) и реагировать на сообщения или события смены времени
 - Получать информацию, правила принятия решений и другие сложные структуры (цели, стратегии) из уровня данных
 - Изменять информацию в уровне данных
 - Реагировать на изменение информации в уровне данных

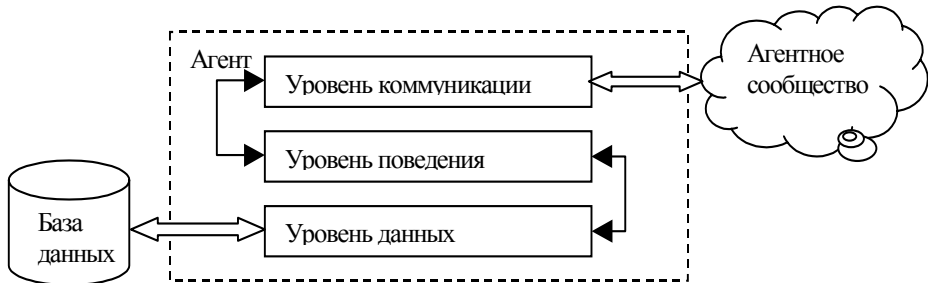


Рисунок 4 – Уровни построения

Заключение

Разработанные средства и принципы построения МАС нашли применение в новой линейке продуктов компании Маджента, в первую очередь в транспортной логистике. На базе инструментальных средств и платформы J2EE были разработаны первые версии системы поддержки операционного планирования транспортных операций и системы моделирования транспортных сетей.

Разработанные инструментальные средства позволяют создавать системы масштаба предприятия, обеспечивая следующие свойства:

- Масштабирование - достигается путем увеличения вычислительных машин в многомашиной среде, а также использования баз данных для хранения промежуточных результатов расчёта;
- Надежность - достигается путем хранения онтологий, сцен и промежуточных результатов расчёта в базе данных, что гарантирует сохранность информации при отказе одного из серверов;
- Безопасность - реализуется с помощью стандарта JAAS, а также уровне доступа к базе данных,
- Кроссплатформенность - достигается использование языка Java и отказом от использования native кода. Поддержка Microsoft Windows и Linux.
- Интеграция и совместимость: все важнейшие информационные составляющие систем (онтологии, сцены, контексты) поддерживают формат XML.
- Распределенность - использование технологий Remote Method Invocation (RMI), Java Servlet, Java Message Service (JMS) для обеспечения потока данных и управления в межмашинной среде.

В ходе разработки были произведены тесты на производительность ОМТ (см. таблицу 1).

В настоящее время проводятся эксперименты и начата разработка многомашиных систем с поддержкой параллельных вычислений задач. Также ведутся работы по созданию унифици-

рованного визуального интерфейса для отображения агентных сообществ, процессов переговоров и принятия решений. Объектная модель унифицированного интерфейса должна обладать свойством модульности и позволять разрабатывать расширения для создания специфических способов визуализации.

Таблица 1 – Результаты испытаний

Число классов	10	25	50	100
Число атрибутов	100	250	500	1000
Импорт онтологии из XML, сек	0,01	0,01	0,03	0,08
Экспорт онтологии в XML, сек	0,004	0,007	0,01	0,02
Компиляция онтологии, сек	0,04	0,05	0,08	0,14
Загрузка онтологии из базы данных, сек	0,11	0,14	0,28	0,64

Список литературы

- [1] П.О. Скобелев. Теоретические основы создания открытых МАС для оперативной обработки информации в процессах принятия решений // Труды 5-й Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-21 июня 2003. – Самара: СНЦ РАН, 2003, стр. 295 - 303.
- [2] В. Андреев, К. Ивкушкин, И. Минаков, Г. Ржевский, А. Сафронов, П. Скобелев. Основные компоненты внутреннего устройства мультиагентной системы // Труды 5-й Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-21 июня 2003. – Самара: СНЦ РАН, 2003, стр. 304 - 316.
- [3] В.А. Виттих, Г.А. Ржевский, П.О.Скобелев. Мультиагентные модели взаимодействия в процессах принятия решений // Труды 4-й Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-24 июня 2002. – Самара: СНЦ РАН, 2002
- [4] В.А. Виттих, Д.В. Волхонцев, Е.А. Гриценко, Г.Д. Светкина, П.О. Скобелев. Разработка мультиагентной системы для моделирования процессов взаимодействия населения и органов исполнительной власти в социальной сфере // Труды 4-й Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-24 июня 2002. – Самара: СНЦ РАН, 2002

АНАЛИЗ ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ МУЛЬТИАГЕНТНЫХ ТЕХНОЛОГИЙ В ЗАДАЧАХ ТРАНСПОРТНОЙ ЛОГИСТИКИ

С.В. Батищев¹, К.В. Ивкушкин¹, Т.В. Искварина¹, В.А. Никифоров², П.О. Скобелев¹

¹ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко 1а
nikiphorov@magenta-technology.ru
тел: +7 (8462) 70-66-84; 70-66-85

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: мультиагентные системы, транспортная логистика, планирование расписаний, эвристики, батч-режим

Abstract

This article describes multi-agent approach for building schedules for logistics systems. The approach differs from the traditional methods of scheduling in its ability of dynamically handling schedule disturbances and modifications that normally happen all the time. In this approach the schedule is represented in a way of a balanced network where every conflict coming in causes a wave of changes that lead the network into another balanced state. It helps to build logistics planners for on-line and batch applications which can flexibly and dynamically react to changes in the environment.

Введение

В настоящее время задача планирования перевозок является актуальной и значимой задачей для многих современных компаний, действующих в условиях неопределенности и высокой динамики изменений спроса и предложения.

Сложность решения этой задачи связана с большим объемом и разнообразием предметных знаний, которые надо учесть при решении задачи, и отсутствием методов решения задач, учитывающих особенности современного бизнеса.

Предлагаемый в данной работе мультиагентный подход позволяет описать предметную область в модели существенно более подробно за счет того, что описание производится на уровне каждого отдельного объекта. Такой подход позволяет более гибко и эффективно описать модель транспортной системы, поскольку, во-первых, появляется возможность достаточно подробно описать как структуру объекта, так и логику его поведения, а во-вторых, логика отделена от программного кода и может быть легко изменена [1].

В настоящей работе мы проанализируем различные области возможного применения мультиагентного подхода на примере задач транспортной логистики с целью выявления сильных и слабых сторон мультиагентного подхода как средства решения транспортных задач.

1 Обзор основных проблем планирования транспортных перевозок

В настоящее время известен ряд компаний, которые обеспечивают решение задачи планирования в логистике: Paragon, CAPS, Manugistics. В основе этих систем лежит традиционные математические алгоритмы, базирующиеся на методах динамического программирования и constraint-программирования [2].

Вместе с тем, для этих систем характерен набор проблем. В первую очередь, данные системы имеют встроенную «жесткую» модель логистических операций (например, развоз грузов из центра), которую трудно или невозможно модифицировать или дополнять. Например, при планировании расписания для перевозок составитель должен учесть множество ограничений и требований, включая: законодательные ограничения для рабочего дня водителя; требования к последовательности загрузки-разгрузки; совместимость грузов, грузовиков и складов; рабочее время складов и многое другое. Учет всех этих ограничений, зачастую нечетко заданных, делает задачу достаточно трудоемкой и ограничивает применимость к ней классических математических алгоритмов.

Во-вторых, системы рассчитаны исключительно на батч-режим, в котором система накапливает заказы, которые в определенный момент (например, в 8:00 утра), вводятся в систему и распределяются по транспортным средствам, что затрудняет применение расписания в реальном времени или учет непредвиденных событий. Например, поломка, опоздание, отзыв принятого заказа и т.д. Особым случаем являются встречные (backhaul) заказы, позволяющие ощутимо увеличить доходы от перевозок за счет подхвата заказов на порожних (чаще всего обратных) рейсах.

В-третьих, расписание, построенное в данных системах, требует обязательной ручной доработки, что представляет собой весьма сложную задачу. Для внесения даже небольших изменений в расписание, необходимо изменять исходные данные и заново с самого начала запускать программу, что может требовать несколько часов.

В-четвертых, в упомянутых системах не учитываются индивидуальные особенности (стратегии, ограничения, предпочтения) всех транспортных средств флота, особенностей клиентов и реализуемых заказов и т.д.

В-пятых, не реализуются «человеческие эвристики» – подходы к построению расписаний, основанные на житейском опыте оператора. Как отдельный момент стоит выделить невозможность машинного преодоления ограничений модели, которые можно было бы обойти в силу их необязательности. Например, время отгрузки может быть сдвинуто по договоренности с оператором, грузовики могут обмениваться фурами для ускорения процесса погрузки, в грузовик часто можно загрузить немного больше заявленной емкости. Поскольку при построении расписания традиционными методами не сохраняется информация о непреодоленных ограничениях, планировщик не возвращается к отвергнутым вариантам даже в тех случаях, если нарушение ограничения было минимально.

2 Применение мультиагентного подхода

2.1 Построение первоначального расписания

Традиционно в классическом подходе построение расписания «с нуля» предполагает процесс последовательного планирования, вырождающегося в перебор вариантов в определенном направлении. Однако в сложных моделях транспортной логистики количество расписаний очень велико, и для того, чтобы построить расписание в разумное время, необходимо ограничить глубину поиска, в результате чего полученное расписание оказывается неэффективным. Полный перебор вариантов в таких условиях невозможен, поэтому большинство систем использует подход последовательного применения ограничений. Например, в первую очередь планируется оптимальное распределение направлений перевозок. Затем на полученные варианты распределяются по существующим грузовикам, и часть вариантов отбрасывается после наложения ограничений совместимости и грузоподъемности. И, наконец, на последней стадии система планирует расписание для водителей. Такой способ, хоть и приемлем, все же малоэффективен, поскольку подразумевает перебор очень большого количества вариантов, которые в конечном итоге отсеиваются в результате наложения ограничений. Полученное расписание

обязательно требует доработки, причем в первую очередь – на предмет обхода ограничений, которые были объявлены в математической модели как обязательные.

В мультиагентной модели каждому объекту ставится в соответствие программный агент, который представляет интересы данного объекта [3]. Каждый объект, будь то груз, склад или грузовик, стремится к определенной цели, и обладает определенными стратегиями, описывающими, как этой цели можно достичь. Например, груз имеет цель быть перевезенным точно в срок. Если есть грузовик, который его перевезет – проблемы нет. Но в ситуации, когда нет ни одного свободного грузовика, который мог бы его взять, груз должен искать варианты решения этой проблемы. Одна из стратегий, позволяющих данную проблему решить – это переговоры с грузами, которые могли бы «подвинуться» во времени и уступить данному грузу свое место. Другая стратегия предполагает вести переговоры с грузовиками, которые могли бы специально скорректировать свой путь так, чтобы доставить данный груз. Грузовик, в свою очередь, стремится к максимальному использованию своей грузоподъемности, и, возможно, согласится на такую сделку.

Как мы видим, в мультиагентной системе каждый объект стремится достичь поставленной перед ним цели, и использует различные стратегии для достижения этой цели. В качестве базовых агентов используются на два типа: «заказы» и «ресурсы». Каждый заказ должен найти ресурс, который его исполнит максимально подходящим для него способом, и наоборот, каждый ресурс стремится найти наиболее выгодный для себя заказ. В процессе построения расписания «с нуля» все ресурсы свободны и все ни один заказ не обслужен, поэтому каждый заказ выбирает себе тот ресурс, который ему больше подходит. Однако скоро свободные ресурсы заканчиваются, и заказы начинают конфликтовать. Для разрешения конфликтов заказы и ресурсы вступают в переговоры с целью найти компромиссный вариант. Особо следует отметить, что условия формирования и разрешения конфликтов построены на правилах модели, вследствие чего объединение индивидуальных интересов объектов в конечном итоге отражает интересы всей системы, т.е. формируется расписание, удовлетворяющее условиям системы и, пусть не обязательно оптимальное, но находящееся в состоянии внутренней гармонии, поскольку построено на взаимном сочетании интересов.

Таким образом, планирование не представляет собой последовательный перебор вариантов – поиск идет «со всех сторон», в результате чего полученное расписание не страдает «однобокостью». Кроме того, такой подход позволяет на уровне стратегий поддерживать «человеческие» эвристики, которые невозможно учесть при последовательном переборе вариантов:

- 1) *Начинать распределение ресурсов следует с тех, которые являются наиболее требовательными в отношении совместимости.* Например, если известно, что замороженный груз может быть размещен только на грузовике, оснащенный холодильником, то, следовательно, грузовики с холодильником являются более ценным ресурсом, и следует в первую очередь рассмотреть варианты целесообразного наполнения этих ресурсов.
- 2) *При планировании маршрута грузовика необходимо учитывать возможность того, что в силу непредвиденных обстоятельств ему придется возвращаться на базу, не закончив маршрута.* Исходя из этого, разумнее всего планировать маршрут так, чтобы первая точка, которую посетит грузовик, была наиболее удаленной от базы, чтобы после нее грузовик естественным путем возвращался на базу.
- 3) *Предпочитать варианты планирования заказов на грузовики, для которых уже запланированы операции погрузки/разгрузки в точках, близких к точкам погрузки/разгрузки данного заказа.* Такие варианты должны «выигрывать» даже если они более дороги в реализации, чем вариант планирования на новый грузовик.
- 4) *Не вовлекать в операцию обмена грузом те грузовики, загрузка которых уже близка к полной.* Риск разрушить уже устоявшийся набор грузов на перевозку достаточно велик и не оправдывает преимуществ операции обмена, ориентированной на разрешение непредвиденных конфликтных ситуаций.

Таким образом, уже на этапе батч-режима мультиагентные системы помогают составлять более реалистичные расписания за счет учета стратегий поведения и предпочтений каждого объекта в отдельности.

2.2 Динамическое изменение расписания

Основной проблемой любого, даже идеального расписания, является тот факт, что в подавляющем большинстве случаев оно исполняется с отклонениями от плана. Практика показывает, что в исполнении любого плана всегда происходят сбои, и поэтому даже план, составленный оптимальным образом, но не предусматривающий вероятных изменений, хуже, чем план, не столь оптимальный, но позволяющий внести без большого ущерба некоторые изменения, вызванные непредвиденными обстоятельствами [4].

Помимо естественных причин в виде опозданий и поломок, в индустрии грузоперевозок существует еще одна причина, вызывающая нарушение составленного расписания – это подхват «третьих» грузов с целью уменьшения порожнего пробега грузовиков. Не будучи включенными в исходное расписание, они неминуемо нарушают план передвижения грузов, и каждый подхваченный груз требует коррекции существующего расписания. В классическом подходе все эти операции: и решение о подхвате груза, и перепланирование – производятся вручную.

В мультиагентной модели потенциальные backhaul-грузы также фигурируют как самостоятельные заказы. С точки зрения грузовиков эти заказы менее приоритетны, чем заказы от собственной компании, и при необходимости выбора «или-или» выбор будет сделан в пользу «своего» заказа. Однако если речь идет лишь о величине расходов на исполнение заказов, грузовик скорее рассмотрит вариант, где он проделает чуть больший путь, но выполнит оба заказа, поскольку это более выгодно.

В общем случае любое нарушение расписания – будь то вследствие появления нового заказа или же вследствие опоздания или поломки грузовика – влечет за собой возникновение конфликта между уже существующим расписанием и возникшими обстоятельствами. Причем следует отметить, что данный конфликт носит локальный характер: изменились не общие правила модели, не законы системы, а просто один из ресурсов опоздал или выбыл из игры, и теперь его работу должны взять на себя другие. Точка возникновения конфликта порождает «волну переговоров», которая будет распространяться по расписанию до тех пор, пока конфликт не будет разрешен. Рассмотрим пример такой ситуации: пусть в систему попал «третий» заказ X, который может быть взят грузовиком А, который согласно расписанию должен развезти пять грузов по пяти разным точкам (рисунок 1). Если грузовик подберет груз X, то он не успеет развезти все пять грузов и вернуться на базу, поскольку закон ограничивает время поездки.

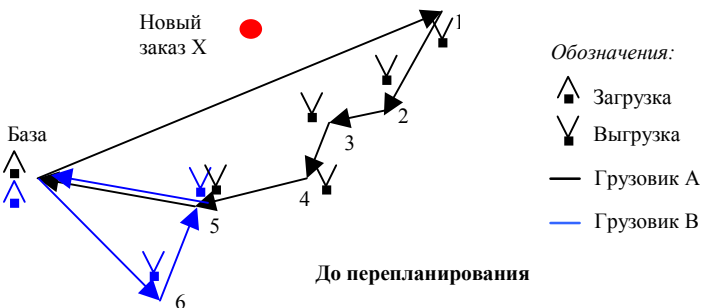


Рисунок 1 – Расписание до принятия нового заказа

Чтобы разрешить возникший конфликт, «волна переговоров» идет дальше и вовлекает в себя грузовик В, маршрут которого проходит через точку выгрузки груза номер 5. В результате

переговоров грузовики достигают следующего компромисса: грузовик А сбрасывает грузы 4 и 5 в точке 4 и кратчайшим путем направляется на базу, а грузовик В совершает небольшое отклонение от маршрута и подбирает в точке 4 груз, предназначенный для точки 5 (рисунок 2). Если такая договоренность не нарушает планов грузовика В, на этом волна переговоров затухает.

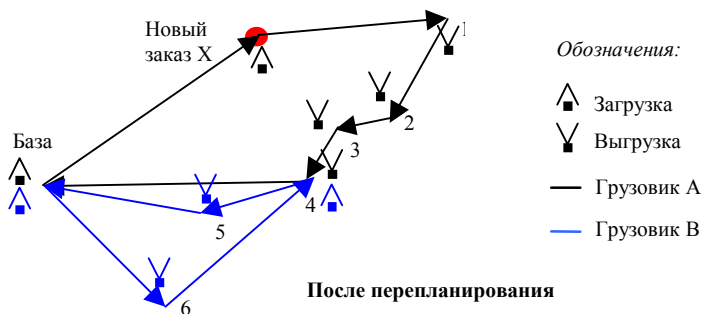


Рисунок 2 – Измененное расписание после принятия нового заказа

Таким образом, в рассмотренном примере существующее расписание было динамически модифицировано так, что в нем появилось место для выгодного заказа.

Заключение

В настоящей работе рассмотрены проблемы построения транспортных расписаний и проведен анализ возможности применения мультиагентного подхода в данной области.

К основным преимуществам использования мультиагентного подхода можно отнести следующие:

- сбалансированный подход к процессу построения расписания;
- учет индивидуальных задач на уровне каждого отдельно взятого заказа и ресурса;
- использование в процессе планирования стратегий, интуитивно используемых человеком на базе опыта;
- возможность эффективной и быстрой коррекции расписания в случае нарушений.

Перспективы развития подхода связываются с поддержкой вторичной логистики, расширением составляющих производственной логистикой, самообучением агентов в ходе работы и способностью агентов принимать стратегические решения.

Список литературы

- [1] В.А.Виттих, П.О.Скобелев. Мультиагентные модели взаимодействия для построения сетей потребностей и возможностей в открытых системах // Автоматика и телемеханика, №1, 2003, стр. 177-185.
- [2] V.Andrejev, S.Batishchev, K.Ivkushkin, I.Minakov, G.Rzevski, A.Safronov, P.Skobelev. MagentA Multi-Agent Engines for Decision Making Support // International Conference on Advanced Infrastructure for Electronic Business, Science, Education and Medicine on the Internet (ISBN 88-85280-63-3), 29 July – 4 August 2002, L'Aquila, Italy, pp. 64-76.
- [3] Batishev S.V., Ivkushkin C.V., Minakov I.A., Rzevski G.A., Skobelev P.O. MagentA Multi-Agent Systems: Engines, Ontologies and Applications // Proc. of the 3rd Intern. Workshop on Computer Science and Information Technologies CSIT'2001, Ufa, Russia, 21-26 September, 2001. – Ufa State Aviation Technical University – Institute JurInfoR-MSU, Vol. 1: Regular Papers, 2001, pp. 73-80.
- [4] Ивкushкин К.В., Минаков И.А., Ржевский Г.А., Скобелев П.О. Мультиагентная система для решения задач логистики // Труды 7-й Национальной конференции по искусственному интеллекту с международным участием ИИ-2000, 24-27 октября 2000, Переславль-Залесский, Россия – М.: Физматлит, 2000, том 2, стр. 789-798.

ПОСТРОЕНИЕ ОНТОЛОГИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ ТРАНСПОРТНОЙ ЛОГИСТИКИ

Д.В. Абрамов¹, В.В. Андреев¹, С.В. Батищев¹, Т.В. Искварина¹, Д.В. Карягин¹, А.В. Сафронов²

¹ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко, 1а
iskvarina@magenta-technology.ru
тел: +7 (8462) 70-66-84; 70-66-85

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: мультиагентный планировщик, базовая онтология, расширенная онтология, транспортная логистика

Abstract

This article describes the approach of building ontologies for multiagent application in transportation logistic. This approach differs from the traditional method of ontology building by extraction of the base level of transportation ontology. The use of such approach allows to specify only ontology broadening specific for particular problem domain area. The article describes examples of a big British truck company ontology and ontology of scheduling of oil transportation by very large crude oil carriers (VLCC).

Введение

В настоящее время все более актуальной становится задача составления расписания для систем транспортной логистики. Перевозки груза по морю, по суше, по воздуху имеют общие принципы, но в тоже время обладают достаточным количеством характерных особенностей, которые следует учитывать при планировании. Сложность построения таких расписаний заключается в учете ограничений и бизнес-правил, специфичных для конкретной предметной области.

Отделение знаний предметной области от алгоритмов планирования позволяет реализовать универсальный инструмент для решения таких задач. Знания предметной области предлагается вынести в онтологию транспортной логистики. При этом расписание удобно составлять с помощью мультиагентного планировщика [1], использующего знание из онтологии.

Естественно, что для решения всего многообразия транспортных задач не может быть построена единственная онтология. Построение такой универсальной, всеобъемлющей онтологии привело бы к потере деталей и ограничений многих конечных транспортных задач. Исходя из этой посылки, предлагается разделить онтологию транспортной логистики на два уровня – базовый и расширенный [2], или уровень прикладных онтологий.

Базовый уровень позволяет задать основные, универсальные понятия (заказ на перевозку, ресурс). На концепты базовой онтологии опираются алгоритмы планирования, а в частности, агенты виртуального рынка транспортной логистики в процессе переговоров по поиску варианта расписания.

В то время как базовая онтология едина для всех транспортных задач, прикладные онтологии специфичны для отдельно-взятого бизнеса. Такие онтологии имеют динамический характер и позволяют развивать онтологии, вносить в них новые понятия, ограничения и правила

одновременно со стороны многих пользователей без изменения структуры системы планирования.

Такое разделение онтологии предметной области для мультиагентных систем обеспечивает «прозрачность» ее построения и позволяет изначально задать правила рассуждения и поведения агентов в более общем виде, не требующем постоянной модификации, независимо от дальнейшего развития динамической части онтологии.

1 Базовая онтология для задач транспортной логистики

При анализе примеров транспортных задач [3], [4] были выделены ряд общих понятий и отношений между ними для данной предметной области. На основе этих понятий была построена базовая онтология для планирования расписаний в задачах транспортной логистики. Рассмотрим подробнее эту онтологию.

Неотъемлемой частью любой транспортной задачи являются понятия заказа (заказа на перевозку) и ресурса, который осуществляет грузоперевозку (например, транспортного средства).

Понятие ресурса представлено в онтологии классом *Resource*, для которого заданы следующие обязательные атрибуты:

- *OptimalSpeed* – оптимальная скорость перемещения.
- *Capacity* – вместимость ресурса
- *InitialLocation* – географическая точка, в которой ресурс входит в систему
- *InitialDate* – дата и время, когда ресурс входит в систему
- *HomeBase* – база ресурса (географическая точка, к которой ресурс приписан)
- *NeedReturnToBase* – булевый атрибут, который указывает, должен ли конкретный экземпляр ресурса возвращаться на базу в конце пути.

Для задач транспортной логистики можно выделить четыре уровня детализации заказа:

- *Точка пути или точка исполнения* определяет место и время одной операции в заказе. Точка исполнения характеризуется географическим положением (уникальное имя географической точки), типом операции (погрузка груза, выгрузка груза, любое действие, не связанное с грузом), ее длительностью, а также периодом, в который должно начаться выполнение операции.
- *Груз*. Основные характеристики груза – его тип и объем/вес.
- *Транспортная инструкция* – инструкция на перевозку груза.
- *Заказ* – набор (одна или более) связанных транспортных инструкций.

Для указанных понятий в онтологию были введены следующие классы объектов:

- Класс *Waypoint* (точка исполнения) с атрибутами *Location* (строковый), *OperationType* (перечислимый), *Duration* (целый, измеряется в минутах), *AvailabilityBeginTime* и *AvailabilityEndTime* (дата/время).
- Класс *Cargo* с атрибутом *Volume* (вещественное).
- Класс *Transportation Instruction*, связывающий классы *Waypoint* и *Cargo*.

В логистических системах на самом верхнем уровне заказ – это сложный объект, который описывает пожелания заказчика. Например, в качестве заказа можно рассматривать доставку пользователю компьютера, который состоит из многих компонент: системного блока, монитора, программного обеспечения и т.п. Но для задач транспортировки такой сложный заказ выродится в набор связанных транспортных инструкций по доставке всех компонент от конкретных поставщиков в место сборки компьютера, а также инструкции по доставке компьютера заказчику. Задача деагрегации заказа на составляющие является отдельной задачей, и обычно на вход транспортных планировщиков поступают уже готовые транспортные инструкции. Исходя из этой постановки, понятие заказа было вынесено за рамки данного исследования.

Рассмотрим подробнее понятие «транспортная инструкция».

Транспортная инструкция состоит из набора точек исполнения (одна и более), которые могут быть связаны с грузом. Например, в транспортной инструкции может быть описано, что нужно перевезти груз G в размере x тонн из точки A в точку B. Транспортная инструкция может быть и как более сложной, так и более простой. Она может состоять только из одной точки исполнения, не связанной с грузом. Примером может служить инструкция на доставку пустого ресурса к географической точке для дальнейшего использования (доставка грузовика к фабрике, которая впоследствии в течение дня использует его по своему усмотрению).

Более сложная инструкция формулируется обычно, чтобы задать заказ на перевозку груза из одной географической точки в несколько других (multi-drop) или, наоборот, из нескольких точек в одну (multi-collect). При этом инструкция может требовать перевезти один или более грузов. В случае, если груз является делимым (100 тонн нефти, 20 паллет стирального порошка) можно задать одну точку погрузки (фабрика, поставщик, порт выработки нефти) и несколько точек выгрузки, в каждой из которых будет разгружена доля груза.

Для некоторых предметных областей важно иметь возможность определить несколько грузов в рамках одной транспортной инструкции. В частности такой подход характерен для курьерских служб, когда требуется в одной доставке привезти несколько разнородных предметов (например, письмо и бандероль), подобранных в разных точках.

Для того, чтобы задать в онтологии такую связь грузов и точек, введено отношение *HaveCargo* между классами *Waypoint* и *Cargo* с атрибутом *Quantity*, который определяет какое количество указанного груза нужно загрузить/выгрузить в указанной точке пути.

Для описания транспортной инструкции также важно понятие заказчика – класс *Customer*, каждый экземпляр которого может быть связан отношением *HaveOrder* (иметь заказ) с одним или несколькими экземплярами класса *Transportation Instruction*.

Кроме указанных классов в базовую часть онтологии введены понятия описатель географической точки (класс *Location*) и описатель маршрута (класс *Route*). Выделение таких классов позволяет определить в виртуальном рынке для планирования расписаний транспортных сетей соответствующих агентов, как описано в [1]. Такой подход позволяет описать некоторые типы точек и маршрутов, которые имеют специфические ограничения или правила прохода (такие как порты, склады или маршруты через северные моря), задать отношения между ними и другими понятиями предметной области, описать ограничения. Кроме того, этот подход дает возможность учесть описанные отношения, ограничения и правила при планировании.

Отметим, что конструктор онтологий компании Magenta [5] позволяет пользователю описывать правила для любого класса или экземпляра в точках принятия решения его агента.

2 Онтология перевозок на грузовиках в Великобритании как пример расширения базовой онтологии транспортной логистики

На базе описанной выше базовой онтологии для задач транспортной логистики была реализована онтология перевозок грузов на грузовиках в Великобритании. Данная задача имеет следующие характерные особенности:

- Существует несколько базовых типов грузов и типов грузовиков, которые отличаются в основном вместимостью и ограничениями на перевозимые грузы.
- Грузовики делятся на грузовики из собственного флота компании и сторонних перевозчиков. Собственные грузовики в конце пути должны возвращаться на базу.
- Для некоторых географических точек существуют ограничения на то, какие грузовики могут войти в эту точку (например, в некоторые точки (склады) могут входить грузовики маленького размера).

Для описания заданной предметной области было создано расширение онтологии, в которое были внесены следующие классы объектов:

- Потомки класса *TransportationInstruction*:

- Заказ на груз, требующий охлаждения – класс *ChilledTI*.
- Заказ на обычный груз – класс *AmbientTI*.
- Ресурсы (потомки класса *Resource*):
 - Грузовик – класс *Truck*. Для этого класса уточнен атрибут *OptimalSpeed*: задано значение по умолчанию, равное 40 км/ч.
 - Грузовик с подвижной боковой загрузкой – класс *CurtanSider* (Capacity = 26 паллет)
 - Средний грузовик с задней загрузкой – класс *Box44* (Capacity = 26)
 - Маленький грузовик – класс *Rigid* (Capacity = 8)
 - Грузовик с двухэтажным кузовом – класс *DoubleDecker* (Capacity = 48)
- Географические точки (потомки класса *Location*):
 - *OnlyDDLlocation*
 - *OnlyVansLocation*

На основании этого расширения онтологии для агентов географических точек заданы следующие правила:

Через точки, принадлежащие к классу *OnlyDDLlocation*, могут проезжать только ресурсы класса *DoubleDecker*.

Через точки, принадлежащие к классу *OnlyVansLocation*, могут проезжать только маленькие грузовики (экземпляры класса *Rigid*).

Через все остальные точки могут проезжать любые грузовики.

Для агента охлаждаемого заказа (агента для класса *ChillOrder*) задано правило на проверку совместимости ресурса: такой заказ может быть перевезен только на грузовиках класса *Rigid* или класса *Box44*.

Таким образом, на выбранном базисе можно описать онтологию перевозки грузов на грузовиках. Отметим, что транспортные инструкции в большинстве случаев в рамках данной задачи состоят из одного груза и двух точек исполнения – погрузка и разгрузка. Рассмотренное расширение базовой транспортной онтологии позволило построить расписание грузоперевозок с учетом основных, самых важных ограничений.

3 Расширения базовой онтологии: морские грузоперевозки

Другим примером транспортных задач могут служить морские грузоперевозки.

Морская логистика включает в себя несколько основных типов перевозок: пассажирские перевозки, перевозка сухогрузов (природные ископаемые, материалы, контейнеры) и перевозка нефтепродуктов, сжиженного газа. Несмотря на разницу ограничений, накладываемую каждым конкретным грузом на параметры судна, который его перевозит, все морские перевозки имеют схожие принципы.

На море не бывает пробок, средняя продолжительность поездки судна много выше чем продолжительность поездки по суши или воздушным транспортом. Движение судна продолжается сутки целиком и не предполагает остановок связанных с отдыхом экипажа. Разгрузка и погрузка грузов на судне более продолжительна и обязательно предполагает наличие специфических терминалов (перевозка нефтепродуктов), наличия у причалов разгрузочных механизмов (контейнеровозы) и т.д.

Средней по сложности планирования, но являющейся одной из самых прибыльных, типов перевозок по морю является перевозка нефти супертанкерами [6]. На базе приведенной базовой онтологии была построена расширенная онтология для задачи перевозки нефти супертанкерами.

В качестве класса ресурса данной онтологии выступает класс *Vessel*, потомок базового класса *Resource*. В этом классе атрибуты базового класса интерпретируются следующим образом:

- *HomeBase* – порт приписки данного судна (Home Port). Соответствуют атрибуту приписки ресурса к некоторой географической точке;
- *Capacity* – вместимость судна или суммарный объем груза, который судно может перевезти (Cargo Volume). Суммарный объем складывается из объема всех грузовых баков судна.
- *InitialLocation* и *InitialDate* – порт и время, в котором судно находится в начале планирования.

Атрибут базовой онтологии *OptimalSpeed* не может быть задан простым числовым значением и является в данной расширенной онтологии расчетным. Расчет производится с использованием дополнительного класса *VesselSpeed*, описанного ниже. Рассмотрим набор атрибутов и дополнительных классов, доопределяющих класс Resource до класса Vessel в расширенной онтологии:

- *Owner* – непосредственный владелец данного судна;
- *Operator* – компания, которая осуществляет управление данным судном, может совпадать с владельцем;
- *Year Build* – год постройки судна;
- *LOA (Length Overall)* – длина судна включая все палубные постройки;
- *Beam* – ширина судна;
- *Draft* – осадка судна. Различают два вида осадки: текущую осадку и максимальную осадку;
- *BTCM (Bow To Centre Manifold)* – численная характеристика оборудования швартовки;
- *TPC (Tons Per Centimeter)* – характеристика судна, используется для расчета осадки судна при увеличении его массы;
- *SCNT (Suez Net Tonnage)* – водоизмещение судна при проходе Суэцкого канала;
- *Flag* – флаг страны регистрации судна;
- *Blue Card* – сертификат которым должно обладать судна при заходе в некоторые порты;
- *Vapour Recovery* – параметр судна, определяющий наличие оборудования очистки воздуха грузовых баков.
- *Hull* – специфика корпуса судна. Различаются суда с двойным и одинарным корпусом;
- *Tropical Deadweight, Tropical Fresh Deadweight, Winter Deadweight, Summer Deadweight, Fresh Deadweight* – максимальные водоизмещения судна в различных географических зонах и климатических условиях, являются общепринятыми характеристиками судна.

Данные параметры судна используются для проверки ограничений накладываемых фрахтователем судна или транспортными инструкциями груза. Так, например, могут быть наложены ограничения на флаг страны, где зарегистрировано судно или на порт и страну порта регистрации. Географические точки (порты), которые транспортная инструкция требует посетить, могут накладывать ограничения на корпус судна (например, в США запрещено принимать суда с одинарным корпусом) или на наличие определенного оборудования.

Специфическим ограничением крупной морской логистики является ограничение на осадку и водоизмещение. Данные два параметра тесно связаны и зависят от множества факторов, таких как количества груза на борту судна, веса топлива, веса экипажа, смазочных материалов, изменений климат и плотности воды по пути следования. Все эти зависимости должны быть описаны в онтологии и учтены при составлении расписания.

Кроме вышеперечисленных характеристик, судно обладает набором свойств, которые сложно описать простыми атрибутами. Для их определения требуется ввести дополнительные классы: *Vettings* (описатели сертификатов), *FuelTanks* (топливные баки), *VesselSpeed* (скорость судна). Рассмотрим подробнее эти понятия.

Безопасность перевозки нефтепродуктов является ключевой проблемой в данной отрасли. В связи с этим все суда должны обладать соответствующими сертификатами, разрешающими перевозку того или иного груза. Класс *Vettings* описывает такие сертификаты. Каждый сертификат характеризуется временем начала действия сертификата, временем окончания действия,

компанией выдающий сертификаты, характеристикой важности сертификата, а также возможностью его продления.

Существует несколько типов топлива, которые потребляют современные суда. Обычно супертанкер использует два типа топлива HFO и MDO. В зависимости от типа двигательной установки судно может потреблять оба типа топлива для движения. Тип топлива HFO используется в основном для передвижения, операций по перекачке нефти и балластной воды. Тип топлива MDO также используется для передвижения, но основным использованием данного вида топлива является выработка электрической энергии в генераторах судна. Таким образом, базовая онтология перевозчиков расширяется специфическими параметрами судна, связанными с потреблением топлива:

- MDO Consumption – потребление топлива MDO в день для выработки электрической энергии;
- Extra Fuel Drydocking – затраты топлива HFO на освобождение грузовых баков от инертного газа при постановке судна в сухой док;
- Extra Fuel Loading – затраты топлива HFO на операции погрузки и разгрузки;
- Minimum HFO, Minimum MDO – минимальное количество топлива которое всегда должно быть в баках судна;
- HFO Safety Margin, MDO Safety Margin – процент от количества топлива необходимого для выполнения операции, который добавляется к полному расходу для контроля доступного объема топлива.

Класс Fuel Tanks описывает баки топлива, которыми обладает судно. Параметрами, характеризующими бак для топлива, являются: тип топлива, цена топлива в баке, дата заливки топлива, текущее количество топлива и объем бака. Пока в баке находится хоть какое-то количество топлива, заливать в этот бак топливо того же типа, но отличающееся хоть чем-то запрещено.

Класс Vessel Speed задает скорости судна в балластном и груженом состоянии, а также расход топлива для каждой скорости. Каждый экземпляр класса Vessel Speed обладает следующими параметрами: значение скорости, потребление топлива HFO, потребление топлива MDO, признак, находится ли судно в балластном или в груженом состоянии. Каждое судно имеет крейсерскую скорость в груженом и балластном состоянии, но может изменять скорость, в заданных пределах, уменьшая или увеличивая потребление топлива, влияя, соответственно, на атрибут OptimalSpeed.

В качестве объекта заказа в морской логистике выступает инструкция на транспортировку нефти. В большинстве случаев при перевозке нефти супертанкерами транспортная инструкция имеет несколько точек погрузки и несколько точек выгрузки нефти, что связано с большими объемами перевозимой нефти. Важно отметить, что для морских грузоперевозок наиболее важным с точки зрения транспортной инструкции является окно доступности первой точки исполнения (т.е. период от *AvailabilityBeginTime* до *AvailabilityEndTime*). Для описания этого периода существует специальный термин - *Lausan*.

Для описания груза «нефть» был введен класс Oil, как наследник класса базового Cargo. Кроме стандартных наследуемых атрибутов, класс Oil расширяется дополнительными атрибутами характерными для груза нефти:

- Freight – стоимость перевозки данного груза. Может задаваться несколькими способами: абсолютным значением, процентом от стандартной стоимости перевозки тонны нефти из пункта А в пункт Б, оплата использования судна по дням;
- Vettings – набор сертификатов которые должно иметь судна для перевозки данного груза;
- Bunker After Load – разрешение на осуществление заправки судна после погрузки груза;
- Grade – тип фракции данного груза;
- Hull – требования к корпусу судна, на котором будет перевозиться груз;

- Address Commission, Broker Commission – проценты от стоимости перевозки, который следует перечислить за операции посредничества в портах и за услуги брокеров;
- Trade – стандартный маршрут перевозки нефти из области ее добычи в область переработки;
- Max Vessel Age – максимальный возраст судна, рассматриваемого для перевозки данного груза;
- Quantity – масса перевозимого груза;
- Max Overage – максимальная масса груза, который может быть перевезен поверх оговоренной в контракте массы;
- Max Speed – максимальная скорость судна, с которой может перевозиться данный груз;
- Demurrage Rate – стоимость одного дня простоя судна по вине фрахтователя;
- Denied Vessels – список конкретных судов на которых перевозка данного груза не возможна в силу различных причин;
- Denied Flags – список запрещенных стран регистрации судов;
- Overage rate – процент от стоимости перевозки тонны нефти сверх оговоренного в контракте количества.

Набор географических объектов так же был расширен с учетом специфики задачи. Были добавлены:

- Понятие Trade, или шаблон пути, так как существует множество стандартных перевозок. Все шаблоны маршрутов движения судов задаются между двумя портами (класс Port является наследником класса Location базовой онтологии), иногда требуется указать промежуточные точки (*Via Point*), например, мыс Доброй Надежды.
- Введены понятие зоны, объединения (союзы) и иерархии точек - каждый порт принадлежит стране, страна входит в экономический союз. В то же время порт принадлежит некоторой географической области. Области имеют иерархичную структуру, начиная со всего мира и оканчивая совсем маленькими объединениями портов.

Таким образом, онтологию, описывающую процессы перевозки сырой нефти супертанкерами, можно описать, основываясь на заданной базовой онтологии. Из базовой онтологии были использованы все классы, класс транспортных инструкций использовался без каких либо изменений.

Заключение

Для упрощения построения точных расписаний для задач транспортной логистики с учетом всех ограничений предлагается вынести знание о предметной области в онтологию. При создании таких онтологий предлагается выделять ядро онтологии (базис) и ее изменяемую часть.

В данной статье также описана базовая онтология для задач транспортной логистики и приведены примеры расширенных онтологий, специфичных для отдельных задач.

Отметим, что исторически первой была разработана онтология морских грузоперевозок. Описанная базовая онтология была сконструирована, используя онтологию морской логистики как пример, а так же рассматривая некоторые другие задачи. Состоятельность выделенного базиса была проверена при построении онтологии для английской логистической компании, занимающейся перевозками на грузовиках. Данная онтология также была представлена в статье.

Использование предложенного подхода дало следующие преимущества:

- Устойчивость базовой онтологии. Описанная базовая онтология является достаточной для построения логистических приложений, что было подтверждено примерами;
- Быстрое построение прикладных онтологий. Выделение базового уровня позволяет описывать только специфичные для каждой конкретной предметной области данные;

- Реализация планировщика расписаний осуществляется один раз для базовой онтологии. Перепрограммирование данной функциональности не требуется для расширенных онтологий;
- Упрощение работы инженера по знаниям. При построении онтологии предметной области инженеру по знаниям не требуется создавать онтологию с самого начала. Инженер использует базовую онтологию и достраивает только те части, которые еще не описаны;
- Выявление скрытых знаний. В процессе расширения базовой онтологии инженер по знаниям переосмысливает бизнес-процессы компании, для которой строится онтология. Выявляются ранее не замеченные проблемы, необходимость изменения процессов работы.

Целью следующих исследований является выделение онтологий промежуточного уровня между базовым и расширенным, онтологий уровня приложения. Примерами таких онтологий могут служить расширение представленной онтологии морских грузоперевозок перевозками сыпучих грузов.

Список литературы

- [1] В. Андреев, М. Андреев, С. Батищев, Т. Искварина, В. Никифоров, П. Скобелев. Мультиагентный конструктор и планировщик транспортных сетей. – Наст. сборник.
- [2] С. Батищев, Т. Искварина, П. Скобелев. Построение развивающихся онтологий для открытых мультиагентных систем в сети Интернет. – // Труды 5-ой Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-21 июня 2003. – Самара: СНЦ РАН, 2003, стр. 324 – 332.
- [3] Michael D. Tusiani. The petroleum shipping industry, PennWell Publishing Company, Tulsa, Oklahoma, 1996
- [4] The Handbook of Logistics and Distribution Management, Alan Rushton, John Oxley, Phil Croucher, Rushton, Oxley and Croucher, 2000
- [5] В. Андреев, С. Батищев, Т. Искварина, П. Скобелев. Разработка новой версии инструментария для создания мультиагентных приложений промышленного масштаба. – Наст. сборник.
- [6] Д. Абрамов, К. Ивкушкин, Д. Карягин, А. Олейников, А. Сафронов, П. Швейкин. Мультиагентная система поддержки принятия решений по планированию грузоперевозок в морской логистике. – Наст. сборник.

МУЛЬТИАГЕНТНЫЙ КОНСТРУКТОР И ПЛАНИРОВЩИК ТРАНСПОРТНЫХ СЕТЕЙ

В.В. Андреев¹, М.В. Андреев², С.В. Батищев², Т.В. Искварина², П.О. Скобелев²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко, 1а
bat@magenta-technology.ru
тел: +7 (8462) 70-66-84; 70-66-85

Ключевые слова: планировщик, мультиагентные переговоры

Abstract

Key requirements for transportation logistics planning systems are proposed. Existing product features are analyzed. Architecture and algorithms of multiagent scheduling system based on J2EE architecture are explained.

Введение

Системы планирования и поддержки исполнения составленного плана являются неотъемлемой частью современной информационной системы поддержки работы транспортной компании.

Можно выделить две категории потребителей таких систем:

- *Операторы и диспетчеры:* составляют план перевозок на определенный период, отслеживают исполнение плана, изменяя план, если этого требуют изменившиеся условия среды.
- *Консультанты и аналитики:* проектируют или оптимизируют сеть поставок компании, описывая ограничения предметной области и моделируя планирование/исполнение перевозок.

Операторы и диспетчеры предъявляют следующие требования к системе планирования:

- Планировать поток заказов на множество доступных транспортных средств. Поддерживать при планировании ограничения предметной области (совместимости грузовиков и грузов, совместимости грузовиков и пунктов погрузки/разгрузки, пропускная способность склада и другие).
- После формирования расписания пользователь должен иметь возможность инициировать пересмотр расписания, если решение его не устраивает. Такое часто происходит, если не все ограничения или критерии диспетчера учтены в системе (см. [1]). Например, диспетчер может считать, что целесообразно направить грузовик нужного типа в место скопления потенциальных встречных перевозок (см. [2]). В этом случае система должна предложить ему новый вариант расписания, учтя пожелания пользователя, даже если это ухудшит вариант по формальным критериям системы.
- Возможность ввести новые заказы (или изменить существующие) после составления расписания.
- На фазе исполнения расписания система должна поддерживать «вбрасывание» в расписание новых заказов или событий реального мира (например, опоздание, отмена заказа). При

этом часто вступают в силу целый ряд новых ограничений предметной области (например, по реформированию грузов, уже подготовленных для погрузки).

- Система должна отвечать требованиям решения уровня предприятия (масштабируемость, надежность, безопасность).

Консультанты и менеджеры предъявляют следующие требования к системе проектирования сети:

- Возможность описать понятия, типы отношений и правила принятия решений без перепрограммирования системы.
- Возможность сконструировать сеть компании на основе заданных понятий и правил. Для каждого ресурса могут быть заданы его уникальные свойства и правила.
- Возможность смоделировать поведение сети, вводя в нее различные потоки заказов на перевозку.
- Возможность использовать модель сети в процессе тактического планирования.

1 Архитектура и принципы работы мультиагентного конструктора и планировщика транспортных сетей

Как было показано в [2], для решения задач транспортной логистики применим и эффективен мультиагентный подход. На основе этого подхода компанией Magenta был спроектирован набор компонент для решения задач транспортной логистики (см. рисунок 1).

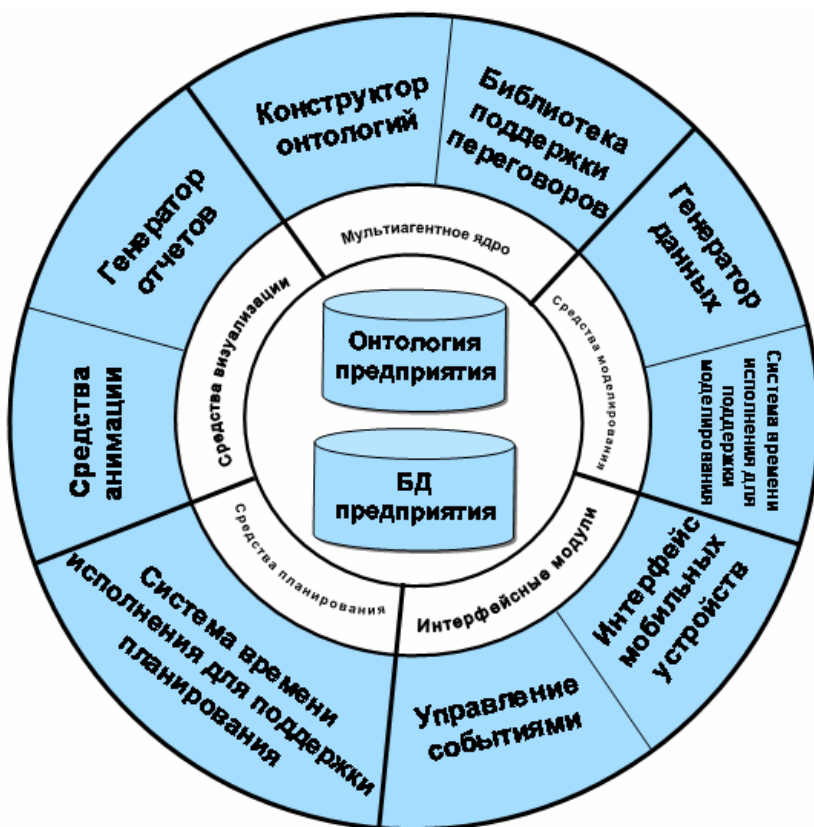


Рисунок 1 – Архитектура мультиагентной системы для планирования перевозок

Разработанные компоненты сгруппированы в два программных пакета, отвечающих требованиям описанных выше групп пользователей:

- Планировщик транспортных сетей. Включает все перечисленные на рисунке модули кроме средств моделирования. Конструктор онтологий планировщика имеет ограниченную функциональность и адаптирован под нужды оператора (небольшие изменения в бизнес-правилах без существенной переработки классов объектов и отношений).
- Конструктор транспортных сетей. Включает все перечисленные на рисунке модули кроме средств управления событиями и интерфейса для мобильных устройств. Конструктор онтологий в этом пакете имеет полную функциональность для описания разнообразных транспортных сетей.

Разработанные компоненты основаны на технологии J2EE и используют реляционную базу данных для хранения состояния агентов, что позволило обеспечить масштабируемость и надежность разработанного решения (см. [3]).

Основным модулем комплекса является система времени исполнения для поддержки планирования. Ее основные компоненты представлены на рисунке 2:

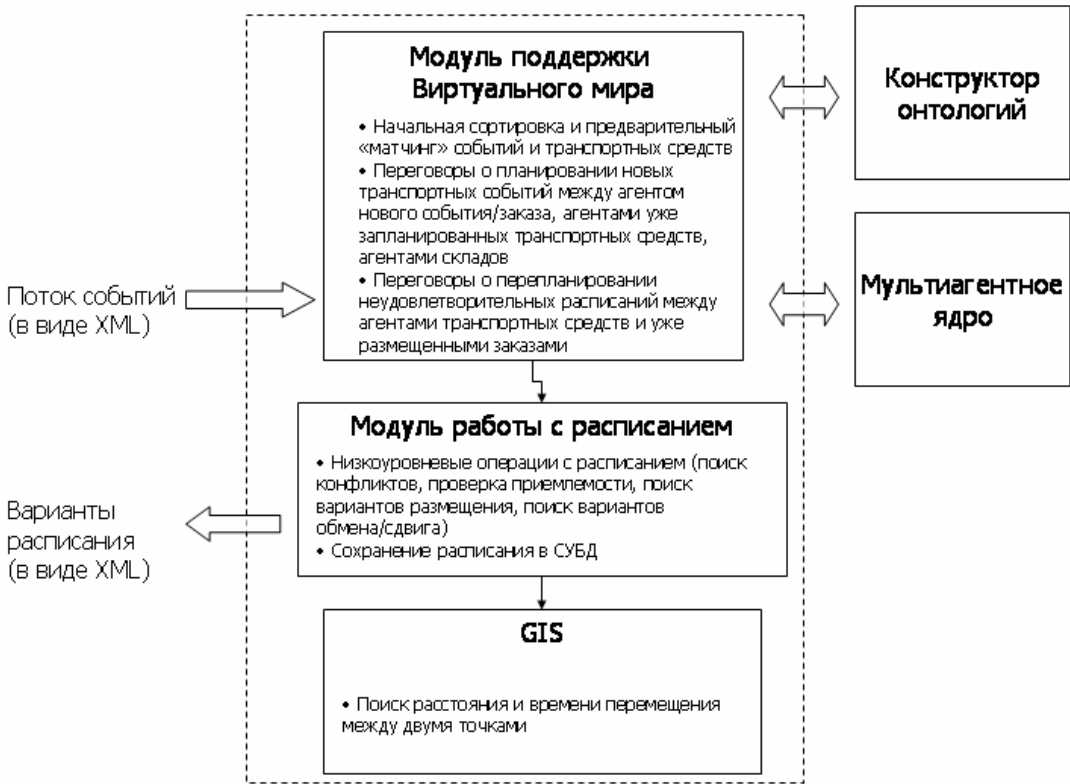


Рисунок 2 – Архитектура система времени исполнения для поддержки планирования

Алгоритм работы системы разбит на несколько этапов, которые представлены на блок-схеме (рисунок 3).

На каждом из этапов агенты действуют в соответствии со своими индивидуальными стратегиями и протоколами переговоров. Так, например, агент компании может иметь следующие стратегии:

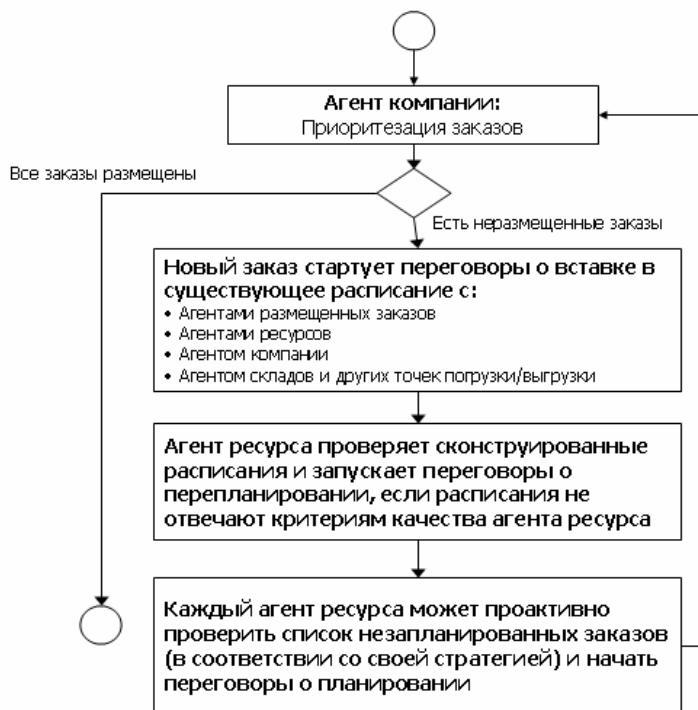


Рисунок 3 – Основные этапы работы системы

- Начинать планирование с наиболее жестко ограниченных заказов (например, с тех, которые могут быть перевезены исключительно на грузовике с холодильным агрегатом).
- Начинать планирование с наиболее дальних заказов (выгодно, если хотим повысить устойчивость расписания к изменениям, так как выполнять перепланирование вблизи базы обычно легче). Альтернатива – начинать планирование с наиболее близких заказов (выгодно, если хотим иметь открытые слоты для встречных перевозок).

2 Протоколы переговоров для решения задач планирования и перепланирования

Рассмотрим основные принципы реализации переговоров в рассматриваемой системе на примере переговоров агента нового заказа с остальными участниками.

Эти переговоры для увеличения производительности были разбиты на несколько этапов (см. рисунок 4):

На каждом из шагов выполняются следующие действия:

- Предварительный «матчинг».
 - Агент заказа проверяет соответствие каждого ресурса своим онтологическим критериям и выполняет их сортировку.
 - Агент заказа запрашивает каждый ресурс, хочет ли он планировать заказ. Ресурс проверяет свои онтологические правила (например, если он загружен на более чем 90 процентов, то может отказаться) и отвечает заказу.

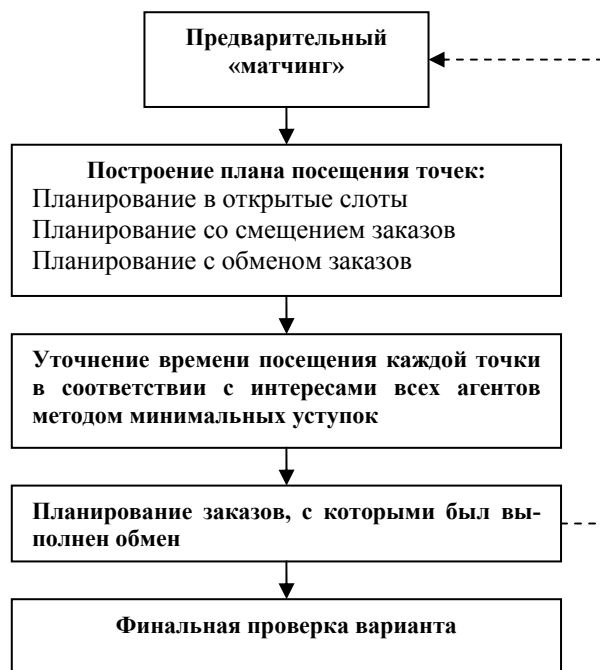


Рисунок 4 – Шаги переговоров агента нового заказа с другими участниками

- Построения плана посещения точек.
 - Агент заказа запрашивает у агента ресурса разрешение на размещение в «открытый слот».
 - Агент ресурса ищет варианты размещения точек погрузки/выгрузки заказа. Агент ресурса проверяет найденный варианты по своим стратегиям (например, отвергать варианты с неэффективной географией) и ограничениям (например, по порядку загрузки/выгрузки) и отбирает лучшие.
 - Агент ресурса возвращает вариант заказу, и тот принимает решение, продолжать ли поиск дальше.
 - Агент заказа запрашивает у агента ресурса разрешение на размещение «со сдвигом» (если на предыдущем шаге решение его не удовлетворило).
 - Агент ресурса запрашивает агентов всех размещенных на нем транспортных инструкций, согласны ли они на сдвиг. Заметим, что имеется возможность сокращения списка сдвигаемых агентов (например, агенты только этого рейса).
 - Агенты размещенных заказов имеют право отказаться от сдвига, если их стратегии/эвристики не разрешают это (например, заказ большого объема менее склонен к сдвигу, чем небольшой заказ).
 - Агент ресурса ищет варианты размещения со сдвигом согласившихся агентов в пределах того же ресурса. Агент ресурса проверяет найденный варианты по своим стратегиям и ограничениям и отбирает лучшие.
 - Агент ресурса возвращает вариант заказу, и тот принимает решение, продолжать ли поиск дальше.
 - Агент заказа запрашивает у агента ресурса разрешение на размещение «с обменом» (если на предыдущем шаге решение его не удовлетворило). Переговоры выполняются аналогично предыдущему этапу. Планирование отличается тем, что если ка-

кой-то из размещенных заказов «мешает» размещению нового, у него запрашивается разрешение переместить конфликтный заказ на новый ресурс (с согласия ресурса и заказа).

- Уточнение времени размещения каждой точки маршрута методом минимальных уступок.
 - Агент заказа предлагает оптимальное время со своей точки зрения (например, стараясь не менять свое старое положение до сдвига или целясь в свое целевое время, если оно задано).
 - Агент ресурса проверяет по своей стратегии, является ли такое размещение приемлемым (типичный критерий при этом – время простоя). Если нет, то ресурс возвращает свой вариант размещения в рамках окна заказа.
 - Агент заказа проверяет по своей стратегии, является ли такое размещение приемлемым, и если нет, то просит ресурс об уступке.
 - Агент ресурса определяет по стратегии, хочет ли он продолжать переговоры и если да, то предлагает шаг на встречу (точку менее выгодную по времени простоя, но более близкую к начальной точке заказа, шаг определяется стратегией).
 - Агент заказа проверяет, удовлетворяет ли его найденное решение и хочет ли он продолжать переговоры (например, по максимальному количеству шагов).

Аналогично переговоры выполняются между агентом ресурса и склада.

Заключение

Описанный алгоритм демонстрирует приемлемые результаты при планировании расписаний «с нуля» и очень эффективен для локального пересмотра расписаний. Это позволяет использовать систему для решения задач, которые ранее требовали ручного планирования.

Безусловно, есть целый ряд направлений для развития системы:

- Поддержка вторичной логистики (водитель, фура и тягач как отдельные дополнительные объекты планирования со своими ограничениями и предпочтениями).
- Поддержка встречных предложение и перекрытия ограничений (загрузить на одну паллету больше, чем положено по нормативной емкости без превышения жесткого ограничения по весу, но сэкономить дополнительный рейс грузовика).
- Поддержка механизмов адаптации и обучения. Кластеризация похожих заказов и ресурсов в группы. Оценка эффективности стратегий и их настройка на основе опыта.

Список литературы

- [1] Бадашкин В.А., Косяков С.В Опыт создания системы планирования грузоперевозок по городу с использованием ГИС / В журнале «Информационный бюллетень. ГИС-Ассоциация» №3,2000, с.61-62.
- [2] С.Батищев, К. Ивкушкин, Т. Искварина, В. Никифоров, П. Скобелев. Анализ возможности и эффективности применения мультиагентных технологий в динамических задачах логистики. - Материалы этой конференции.
- [3] В.Андреев, С. Батищев, Т. Искварина, П. Скобелев. Инструментальные средства для разработки мультиагентных систем промышленного масштаба. - Материалы этой конференции.

СОЗДАНИЕ МУЛЬТИАГЕНТНОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО ПЛАНИРОВАНИЮ ГРУЗОПЕРЕВОЗОК В МОРСКОЙ ЛОГИСТИКЕ

Д.В. Абрамов¹, К.В. Ивкушкин¹, Д.В. Карягин², А.В. Олейников¹, А.В. Сафронов², П.К. Швейкин¹

¹ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко, 1а
karyagin@magenta-technology.ru
тел: +7 (8462) 70-66-84; 70-66-85

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: *нефтеперевозка, компания-оператор, планирование размещения заказов, составление расписаний, мультиагентный планировщик, онтология*

Abstract

This article describes system for scheduling of oil transportation by very large crude oil carriers (VLCC) based on multiagent approach and ontology of oil transportation logistics

Введение

Мир современного бизнеса в последние годы приобретает все большую динамику, современное состояние рынка диктует компаниям необходимость более гибко реагировать на потребности рынка и своих клиентов, и все большее значение при этом приобретают средства автоматизации бизнес-процессов. Одной из интересных с точки зрения автоматизации сфер бизнеса являются задачи транспортной логистики, в частности, задачи планирования морских грузоперевозок.

Предметная область морских грузоперевозок в целом является довольно обширной и объединяет в себя перевозки сухих сыпучих грузов, контейнерные перевозки и танкерные перевозки жидких грузов (различных видов топлива).

Для решения задач оперативного планирования была разработана мультиагентная система составления расписаний перевозок грузов, которая и описывается в данной статье.

1 Описание предметной области

С целью упрощения условий при создании системы было решено ограничиться решением задачи планирования морских нефтеперевозок на сверхбольших танкерах водоизмещением 250-300 тысяч тонн. В соответствии с мировой терминологией, их называют VLCC (very large crude carrier).

Назначением разработанной системы является автоматизация задачи планирования размещения заказов и составления расписаний танкеров флота компании-оператора, занимающегося морскими нефтеперевозками.

В ходе исследования предметной области [1] было определено, что количество судов-перевозчиков такого объема является сравнительно небольшим (около 800 судов во всем мире). Кроме того, было определено, что, как правило, перевозки таких объемов нефти производятся только одним заказчиком одновременно, т.е. можно рассматривать ситуацию, что в каждый момент на судне-перевозчике размещается только один заказ.

В ходе анализа предметной области было выяснено, что для решения задачи планирования перевозок уже созданы некоторые программные системы (например, Route99), однако существующие системы имеют ряд недостатков, которые могут быть преодолены путем построения программного планировщика на основе мультиагентных технологий. В частности, среди недостатков традиционных системы можно отметить схему одновременного планирования всех заказов, когда при поступлении нового заказа, или при необходимости перепланирования одного из заказов необходимо проводить операцию планирования для всех заказов. Полное перепланирование не всегда допустимо в данной предметной области, так как заказчик заранее извещается о судне, на которое запланирована перевозка его груза, и составляется контракт, в котором судно-перевозчик может быть точно определено, и потому перепланирование заказов зачастую может являться нежелательной процедурой.

В ходе дальнейшего исследования были определены основные концепты системы, выработаны представления о необходимых компонентах системы, определена ее архитектура.

2 Архитектура системы

Состав и структура объектов предметной области в системе описывается при помощи онтологии. Основным компонентом системы является мультиагентный планировщик, решающий задачи планирования и перепланирования размещения заказов и составления расписаний судов.

База данных системы является хранилищем онтологий, копии данных о текущем состоянии рынка нефтеперевозок (снимок, или snapshot), сценариев (или сцен, scenario или scene). Отдельное место в структуре базы данных занимает блок, хранящий результаты работы модуля планирования.

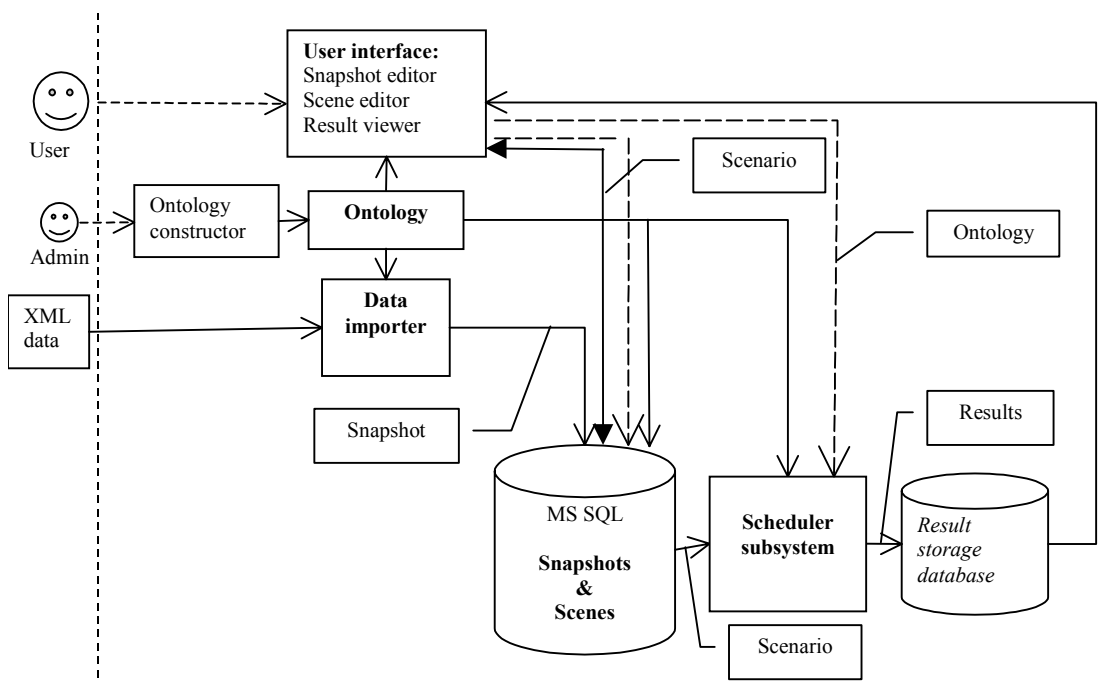


Рисунок 1 – Архитектура системы

Snapshot является набором данных, описывающим текущее состояние системы, хранящим информацию о судах, их расписаниях, заказах, портах, расстояниях между портами, ограниче-

ниях географических объектов, накладываемых на суда и т.д. Структура данных снимка описывается в онтологии. Заполнение снимка данными производится при помощи модуля импорта данных (data importer), использующего в качестве входного формата XML, структура которого генерируется на основе онтологии.

Архитектура программной системы представлена на рисунке 1.

Работа мультиагентного планировщика основывается на обработке событий, являющихся отображением событий реального мира в мире системы. Совокупность снимка, событий и связанных с ними онтологии называется сценой. Результаты планирования сохраняются в отдельную структуру данных в рамках базы данных системы. В результате обработки сцены на выходе планировщик генерирует набор некоторых выходных данных, объединяющих информацию о процессе принятия решений, расписания судов и финансовые показатели, связанные с грузами, судами и флотом в целом.

3 Онтология системы

Онтология системы содержит 3 раздела: описание модели физического мира, описание модели отношений между объектами физического мира, и описание модели поведения агентов, соответствующих этим объектам.

На уровне описания модели физического мира было выделено 11 типов объектов (концептов) и более 250 атрибутов.

Основные онтологические концепты:

- Vessel – танкер, судно-нефтеперевозчик;
- Cargo – груз сырой нефти;
- Company – компания;
- The Operator – компания-оператор;
- Competitor – компания-конкурент;
- Customer – компания-заказчик;
- Location – географические объекты;
- Port - порт;
- Via Point – промежуточная точка;
- Canal - канал;
- Singapore Strait – Сингапурский пролив;
- Suez canal – Суэцкий канал;
- Route – маршрут;
- Region – регион;
- Union – союз (пример – европейский союз);
- Country – страна;
- Area – географический регион;
- Trade – стандартный маршрут нефтеперевозки (например, западная Африка - восточный Берег США);

Каждый из объектов имеет достаточно большое количество атрибутов. Часть атрибутов могут быть вычислимыми на основе значений других атрибутов, формулы для вычисления задаются скриптами.

Кроме того, физическая часть онтологии описывает структуру событий системы, включающих в себя большое количество различных событий, обрабатываемых планировщиком – это группы событий, связанных с появлением нового заказа, события необходимости перепланирования заказа, события опоздания перевозчика и т.д. Всего система поддерживает 21 тип событий. Возможно создание пользовательских типов событий путем наследования от существующих событий для обработки событий, связанных с пользовательскими объектами в систе-

ме. В онтологии также хранится описание структуры нескольких таблиц, хранящих данные, необходимых для работы планировщика. Фрагмент семантической сети, описывающей раздел физической модели мира в онтологии системы, приведен на рисунке 2.

Важной особенностью данной системы является введение стратегий, которые используются агентами объектов виртуального мира для оценки вариантов решений, создаваемых модулем планирования. Стратегии в системе описываются набором скриптов, оценивающих варианты с точки зрения применимости для данного агента. Для каждого варианта расписаний, создаваемого модулем планирования, производится оценка применимости и вычисление значений рейтинга для агента каждого объекта, затронутого в данном варианте, на основе которых строится финальный рейтинг варианта, на основе которого все варианты ранжируются и для дальнейшей обработки планировщиком принимается вариант с наивысшей оценкой. Изменяя скрипты стратегий, и задавая объектам различные стратегии, можно изменять критерии принятия решений планировщиком. В онтологии также можно описать структуру дополнительных таблиц, хранящих данные, необходимые для реализации стратегий.

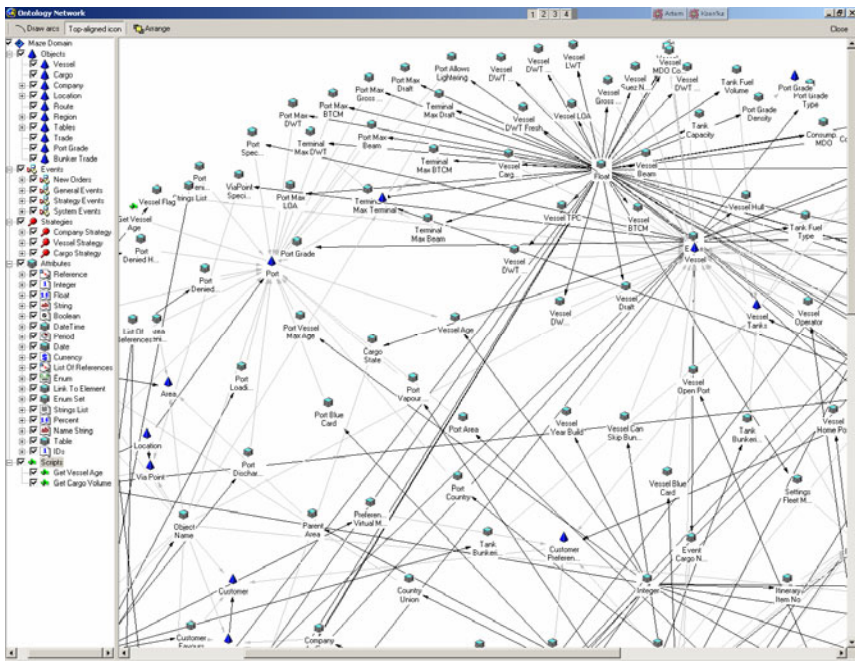


Рисунок 2 – Фрагмент семантической сети онтологии системы

Секция виртуальных отношений описывает агентов ресурсов и заказов, связанных с объектами физического мира и описывает условия возникновения отношений соответствия (matching) между ними. Основная часть ограничений матчинга накладывается на перевозчика географическими объектами, объектом груза, и объектами регионов.

- Отношения матчинга агентов, заданные в системе:
- Cargo with Vessel (между Грузом и Судном)
- Vessel with Port (между Судном и Портом)
- Vessel with Terminal (между Судном и Терминалом)
- Vessel with Union (между Судном и Союзом)
- Vessel with Country (между Судном и Страной)
- Vessel with Area (между Судном и Регионом)

Преимуществом использования онтологии в данной системе является возможность изменения структуры объектов, в результате которого могут быть созданы дочерние типы объектов, новые типы событий, порожденные от базовых, и т.д. Изменяемым параметром также может выступать набор ограничений, накладываемых одними объектами (например, портами), на значения атрибутов других (например, судов). Кроме того, онтология позволяет добиться изменения критериев принятия решений и учета индивидуальных предпочтений агентов системой изменением скриптов стратегий.

Онтология также имеет секцию времени выполнения (Runtime), содержащую описание структуры результатов и отдельных финансовых показателей. При дальнейшей работе над системой планируется введение возможности задания пользовательских показателей, которые планировщик будет рассчитывать для грузов, судов и флота в целом автоматически в ходе планирования.

4 База данных системы

С учетом необходимости хранения в снимке входных данных системы некоторого набора объектов, состав и структура которых может быть изменена с помощью конструктора онтологий, было ясно, что использование традиционной табличной модели данных не является допустимым, в результате чего было принято решение о реализации объектной модели данных средствами реляционной СУБД. Результатом данного решения стало создание базы данных, способной хранить данные о множестве объектов с произвольным составом атрибутов. В качестве СУБД был выбран Microsoft SQL server.

Главной отличительной особенностью используемой модели данных явилось то, что в качестве структурообразующего элемента, определяющего структуру объектов в данной БД, выступает онтология – каждый объект и атрибут, хранимый в базе, имеет ссылку на соответствующий концепт или атрибут концепта в онтологии.

Для полной поддержки прозрачной связи структуры базы данных с онтологией в системе были реализованы средства импорта-экспорта данных и компоненты доступа к объектной модели данных, обеспечивающие прозрачную работу с объектной моделью данных так же, как и с обычной реляционной.

В качестве входного формата данных в системе используется XML. Для генерации структуры XML, соответствующей структуре онтологических концептов был создан модуль генерации структуры XML-файла. для непосредственного импорта данных в объектную базу и связывания XML-данных с онтологией был создан модуль импорта, позволяющий установить соответствие XML-объектов и концептов онтологии. При заполнении снимка (snapshot) данными производится сохранение онтологии в базе данных.

В базе данных также реализованы средства журналирования изменений, позволяющие отследить изменения, произведенные пользователями.

Кроме блока, хранящего данные snapshot, в базе данных созданы еще два блока – один из них предназначен для хранения сценариев, создаваемых пользователем в ходе работы с системой при помощи средств редактирования сцены пользовательского интерфейса. Второй отдельный блок – это блок хранения результатов работы планировщика. Описание данного блока в данной статье будет дано ниже.

5 Визуальный интерфейс системы

Визуальный интерфейс системы объединяет в себе набор средств для работы пользователя со всеми компонентами системы:

- средства просмотра и редактирования онтологий (ontology editor);
- редактор снимков БД (snapshot editor);

- средства редактирования сцены – редактор событий, модуль назначения стратегий агентам отдельных объектов в системе;
 - средства запуска и управления процессом планирования;
 - средства анализа результатов работы модуля планирования, и средства отчетности;
- Иллюстрация интерфейса системы представлена на рисунке 3.

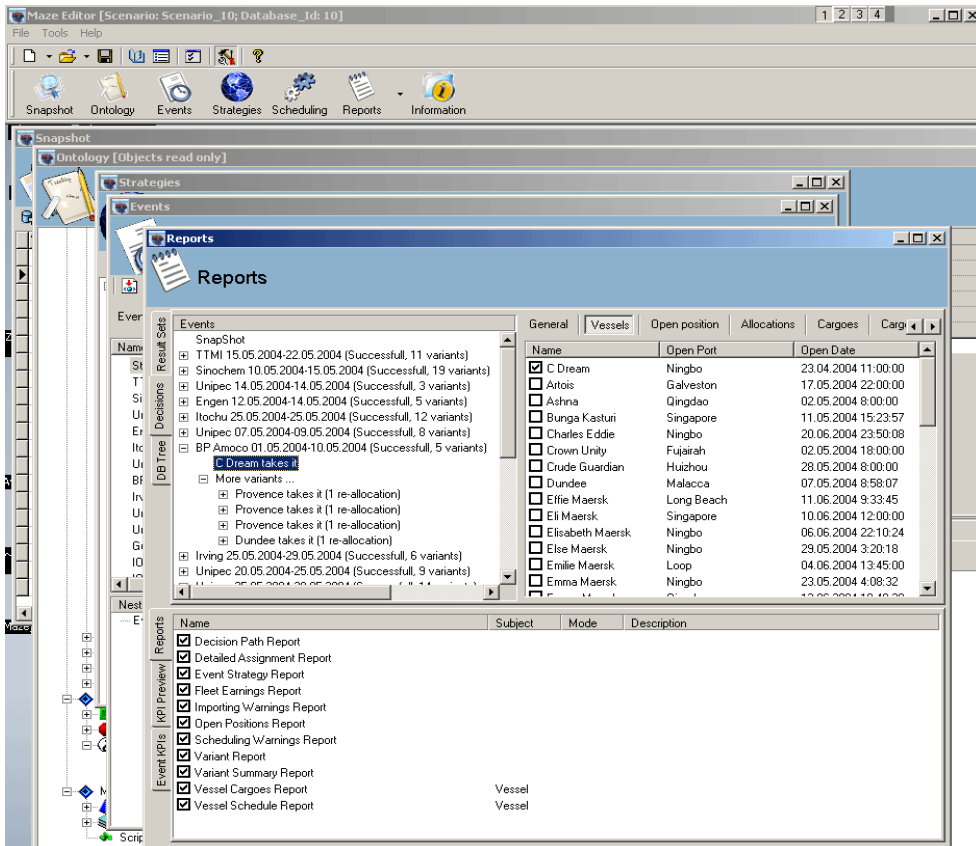


Рисунок 3 – Визуальный интерфейс системы, просмотр результатов

Встроенный редактор онтологий [5] обеспечивает возможность изменения онтологии в рамках ограничений системы – задание дополнительных атрибутов существующим объектам, создание дочерних типов объектов и событий, редактирование условий.

Редактор снимков БД предоставляет пользователю возможность просмотра и редактирования всех входных данных, загруженных из пользовательского XML и используемых планировщиком.

Редактор событий позволяет задавать пользователю события, которые будет обрабатывать планировщик. Часть событий, связанных с появлением в системе грузов, создается автоматически, на этапе создания сценария на основе незапланированных грузов из Snapshot'a. Структура и количество событий является определяемым в онтологии, поэтому редактор событий также является гибким компонентом, использующим для своей работы онтологию.

Модуль назначения стратегий агентам системы позволяет каждому из грузов, судов, и агенту компании-перевозчика назначить индивидуальные стратегии из списка определенных в онтологии.

Интерфейс имеет внутри себя средства запуска и управления процессом планирования (обработку контрольных точек, в ходе которых возможен ручной выбор варианта принятия решений).

Кроме того, интерфейс содержит в себе средства быстрого анализа ситуации, средства просмотра основных результатов планирования, и включает в себя модуль вызова внешних отчетов, реализованных средствами системы создания отчетов Crystal Reports.

6 Мультиагентный планировщик

Одним из основных компонентов системы является мультиагентный планировщик. Данный компонент реализован в виде отдельного консольного приложения, получающего на вход сцены для планирования и выдающего в результате своей работы построенные варианты расписаний. Реализация компонента в виде отдельного приложения обоснована использованием дополнительных библиотек мультиагентного ядра RC Engine 2.0 [4, 6], которые не используются при работе интерфейса, онтологии и модулей работы с базой данных.

Общий алгоритм работы компонента выглядит следующим образом:

- Загрузка исходной сцены;
- Загрузка, планирование, сохранение результатов поступившего события;
- Сохранение общих результатов планирования;
- Завершение работы компонента.

При загрузке исходной сцены, происходит зачитывание и интерпретация онтологии, на которой построена сцена. По этой онтологии загружается схема географии. Создаются агенты терминалов, портов, стран, союзов, регионов. Создаются объекты маршрутов, соединяющих те или иные порты. Данные маршруты являются часто используемыми морскими путями.

После загрузки географии система осуществляет загрузку ресурсов, осуществляющих перевозку. Каждый ресурс-судно имеет большое количество характеристик, параметров. Так, например, для судов загружаются измерения (осадка, ширина, длина), потребление топлива для разных операций (современные суда потребляют несколько типов топлива, например HFO, MDO, GAS), скорости в различных состояниях судна, а также расход каждого топлива при движении на каждой скорости, набор разрешающих сертификатов, начальное местоположение и расписания. В начальном расписании судна могут быть запланированы такие операции как постановка судна в сухой док, offhair (действия, осуществляемые владельцем судна), fixture (ранее запланированный груз, выполнение перевозки которого уже началось или начнется в будущем).

После загрузки сцены и создания всех агентов географии и агентов ресурсов, из базы данных по одному загружаются события. Наиболее распространенным событием является приход обычного spot заказа, являющегося обычной перевозкой груза нефти по задаваемому trade'у. (Trade – стандартная перевозка из региона добычи в регион переработки нефти, обычно в регионах специфицируются конкретные порты, а также указывается путь, по которому должно следовать судно, например, trade: AG/USGviaSC – стандартная перевозка из Арабского залива (Arabian Gulf) в US Gulf через Суэцкий канал (Suez Canal)) Также система поддерживает другие типы событий: планирование постановки судна в сухой док, обработка ситуации поломки судна, опоздания/прихода раньше в указываемый порт, приход и уход нового судна во флот компании оператора, сдача/взятие в аренду судна (TC IN/OUT Orders), а также другие.

Рассмотрим алгоритм планирования основного события – приход нового груза (spot contract). Первым делом составляется список судов, которые теоретически могут перевезти груз данного контракта. (В спецификации контракта может быть задан список разрешенных и запрещенных для использования судов, может быть указано конкретное судно, на котором должна осуществляться перевозка.) На полученный список судов накладываются простые ограни-

чения спецификации контракта, которые не зависят от времени. Среди таких ограничений выступают:

- Возможность судна перевезти объявленный в контракте объем груза;
- Возможность судна зайти в специфицированные порты по своим измерениям (ширина, длина судна, ограничение осадки является ограничением, зависящим от времени);
- Ограничение на флаг порта приписки судна;
- Ограничение на владельца судна;
- Ограничение на корпус судна (суда бывают с одинарным и двойным корпусом);
- Ограничение контракта и портов на возраст судна;
- Ограничение портов на наличие у судна необходимого швартовочного оборудования и терминалов работы с грузом;
- Наличие у судна необходимых сертификатов, обозначенных в контракте.

При необходимости проверку любых ограничений системы можно отключить в онтологии. Также в онтологии можно задать обход ограничения за обозначенную плату.

После проверки ограничений, система приступает к планированию непосредственно расписаний перевозки груза контракта на судах. При построении расписания системой строятся переезды судна по существующим морским путям из порта текущего положения судна через указанную промежуточную точку в порт заправки. По правилам перевозки нефти, заправка судна топливом должна осуществляться до погрузки нефти. Возможно нарушение этого правила, если судно не успевает прийти в первый порт погрузки в указанный в контракте интервал времени, так называемый *Lausan*, если заправка после погрузки разрешена условиями контракта. Далее строятся другие участки расписания по перевозке груза, соединяющие порт заправки с портами погрузки и далее с портами разгрузки. Для каждого порта погрузки/выгрузки может быть задано свое время, в интервал которого надо прийти в данный порт.

Для судов, в расписание которых получилось успешно добавить все необходимые переезды, строится специальная копия всех расписаний флота судов компании оператора – контекст. Компонент построения расписания содержит специальный модуль работы и управления контекстами. В данном модуле осуществляются действия по созданию и поддержке контекстов. (Для оптимального использования памяти, контексты хранятся в виде древовидной структуры, в которой каждый контекст хранит только измененные расписания судов.) После построения контекста, рассчитываются показатели перевозки груза планируемого события. Такими показателями являются стоимость и длительность перевозки, доход судна в день, доход по флоту в день, суммарное время простоя судна во время поездки и т.д.

Основываясь на построенном расписании и рассчитанных показателях, система проверяет оставшиеся ограничения, которые зависят от построенного расписания. Среди таких ограничений можно выделить следующие:

- Ограничение на стоимость перевозки. Стоимость зависит от потраченного топлива во время перемещения или простоя, стоимости погрузки/разгрузки, пошлин за проход каналов и т.д;
- Ограничение на водоизмещение. В зависимости от погруженного количества нефти и топлива изменяется водоизмещение судна. Большинство терминалов обслуживающих портов имеют определенные ограничения на водоизмещение судов, которые они обслуживают. Также во время прохода судна через определенные климатические зона, значение максимального водоизмещения судна меняется, и данное значение не может быть превышено текущим водоизмещением судна;
- Ограничение на осадку. В зависимости от водоизмещения (фактически от количества нефти на борту) изменяется осадка судна.

Если все проверки ограничений проходят успешно, то найденный вариант перевозки удовлетворяет всем требованиям по перевозке данного груза на данном судне. После этого

корректно построенный вариант проходит проверку и сравнительную оценку стратегиями поведения системы. Первой стратегией является допустимость построенного варианта, данную оценку производит груз, судно и компания. После этого каждый из участвующих объектов (груз, судно, компания) проставляют свою сравнительную оценку построенного варианта. По данным оценкам варианты сортируются и лучший из вариантов принимается системой к исполнению. Стратегии задаются в онтологии для каждого типа агента или в сцене для каждого агента в виде специализированных скриптов, написанных на паскале подобном языке.

Кроме обычного способа построения расписания, в морской логистике перевозок нефти существует практика переброски грузов с одного судна на другое. Такая переброска груза осуществляется в случаях, когда груз вновь пришедшего контракта (груз А) может быть перевезен только тем судно, на расписание которого уже запланирована некоторая перевозка другого груза (груз Б). Ситуация разрешается освобождением необходимого судна от груза А, планирование на данное судно груза Б, и перепланирование груза А на любое другое судно (разумеется если это возможно). По схожему алгоритму возможен обмен грузами тремя и более судов. Рассматриваемая система позволяет совершать проверку вариантов обмена до пяти грузов. Отметим, что для поиска варианта обмена тремя грузами человек затрачивает несколько часов, система строит вариант за считанные минуты.

7 Результаты планирования, средства просмотра результатов

В ходе работы модуль планирования проводит проверку ограничений, анализ доступности судов, находит различные варианты размещения грузов. Данные о результатах проверки ограничений, о доступности судов для перевозки, оценки вариантов расписаний сохраняются в базу данных для последующего анализа и оценки. Кроме того, для каждого варианта расписания проводится детальный расчет времени и стоимости операций, с разделением по грузам, судам и суммарные данные по всему флоту в целом. Система сохраняет эти данные для каждого варианта. Полный набор данных, соответствующих некоторому запланированному варианту, называется контекстом. Поскольку полный объем данных, которые необходимо сохранить, достаточно большой, в системе используется дифференциальный метод сохранения. Каждый из контекстов имеет ссылку на предыдущий, родительский контекст, и данные об изменениях, произошедших в данном контексте.

На основе данных, хранимых в базе, в системе строится набор отчетов, позволяющих провести анализ принятия решений, посмотреть финансовые показатели вариантов, расписания отдельных судов, детали планирования перевозок.

Заключение

Результаты пробной эксплуатации системы для планирования размещения реальных заказов на судах крупной компании-оператора, показали, что система является востребованным решением для автоматизации процесса планирования и перепланирования, поскольку позволяет повысить эффективность работы экспертов, занимающихся составлением и оптимизацией расписаний танкеров, предлагая несколько допустимых вариантов расписания, тем самым помогая в процессе принятия решений. Ярким примером является попытка перепланирование груза с использованием перепланирования грузов на 3-4 судах. При ручном планировании на обнаружение и расчет таких вариантов в настоящее время уходит порядка 4-5 часов, время, затрачиваемое системой для просчета таких вариантов – менее 1 минуты.

Таким образом, использование системы может помочь в обнаружении более эффективных вариантов размещения грузов, что в результате приводит к увеличению суммарных доходов компании-оператора и повышает эффективность использования флота. А гибкая система событий, мультиагентное ядро переговоров и онтология, позволяющая изменять критерии приня-

тия решений агентами, делают систему гибким и адаптивным инструментом для поддержки планирования.

Список литературы

- [1] Michael D. Tusiani. The petroleum shipping industry, PennWell Publishing Company, Tulsa, Oklahoma, 1996
- [2] Ивкушкин К.В., Минаков И.А., Ржевский Г.А., Скобелев П.О. Мультиагентная система для решения задач логистики // Труды 7-ой Национальной конференции по искусственному интеллекту с международным участием ИИ-2000, 24-27 октября 2000, Переславль-Залесский, Россия – М.: Физматлит, 2000, том 2, стр. 789-798.
- [3] V.Andrejev, S.Batishchev, K.Ivkushkin, I.Minakov, G.Rzevski, A.Safronov, P.Skobelev. MagentA Multi-Agent Engines for Decision Making Support // International Conference on Advanced Infrastructure for Electronic Business, Science, Education and Medicine on the Internet (ISBN 88-85280-63-3), 29 July – 4 August 2002, L’Aquila, Italy, pp. 64-76.
- [4] Андреев В.В., Ивкушкин К.В., Минаков И.А., Ржевский Г.А., Скобелев П.О. Конструктор онтологий для разработки мультиагентных систем. // Труды 3-ей Международной конференции по проблемам управления и моделирования сложных систем, Самара, 4-9 сентября 2001. – Самара: СНЦ РАН, 2001, стр. 480-488.
- [5] Ivkushkin K., Minakov I., Rzevski G., Skobelev P. “MA DAE: MagentA Multi-Agent Desktop Application Engine” // Proceedings of the 3rd International Workshop on Computer Science and Information Technologies CSIT’2001, Ufa, Russia, 21-26 September, 2001. – Ufa State Aviation Technical University – Institute JurInfoR-MSU, Vol. 1: Regular Papers, 2001, pp. 81-89.

СОЗДАНИЕ МУЛЬТИГЕНТНОЙ СИСТЕМЫ АВТОМАТИЧЕСКОЙ ОБРАБОТКИ, ПРЕОБРАЗОВАНИЯ И КОРРЕКЦИИ ЛОГИСТИЧЕСКИХ СООБЩЕНИЙ СТАНДАРТНЫХ ФОРМАТОВ ОБМЕНА БИЗНЕС-ДАНЫМИ

А.В. Алексеев¹, С.И. Вольман¹, И.А. Минаков¹, А.Д. Орлов¹, М.С. Томин²

¹ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко, 1а
minakov@magenta-technology.ru
тел: +7 (8462) 70-66-84; 70-66-85

²Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: формат обмена бизнес сообщениями, онтология логистики, ANSI X12, EDIFACT, интеграция форматов

Abstract

The article describes new approach for integration of business data interchange formats, based on the ontology of the problem domain. Algorithms for handling of mapping variations and data cleansing support are also introduced.

Введение

Для эффективного управления логистические компании используют сетевые инфраструктуры и информационные системы, включающие базы данных, системы управления бизнес-процессом, ERP-системы. Управление перечисленными информационными системами, как правило, осложняется тем, что они работают с документами в различных форматах. Еще более усложняется ситуация при попытке наладить и автоматизировать совместный бизнес между партнерскими компаниями, когда приходится налаживать взаимосвязи как между различными информационными службами компаний-партнеров, так и со сторонними организациями. В результате отделы и службы компаний образуют десятки разнообразных информационных и организационных связей.

И хотя уже сейчас использование стандартных форматов обмена бизнес-данными очень широко распространено, поскольку существенно снижает затраты на взаимодействие партнеров по бизнесу, сам процесс интеграции является сложным, долгим и крайне болезненным делом.

Как правило, на сегодняшний день проблема перевода сообщений между различными форматами решается путем ручного конструирования схемы преобразования форматов с помощью некоторой программы интеграции приложений (например, BizTalk). Сложность такого решения состоит в том, что требуется серьезный предварительный экспертный анализ, помогающий выяснить семантические соответствия полей различных форматов. В случае же, если формат был адаптирован под нужды фирмы, или в рамках стандартного формата происходит интерпретация полей, специфичная только для данного клиента, процесс выявления соответствий может быть очень сложен и долгов. На текущий день, по оценкам экспертов, построение соответствия одного раздела формата в рамках пакета стандартов занимает около недели. На то же, чтобы полностью интегрировать информационные службы двух компаний, уходит не менее полугода, что является очень дорогостоящим решением, к тому же не отвечающим предъявляемым рынком требованиям к динамике и скорости реакции.

Предлагаемый подход позволяет автоматизировать процесс регистрации новых форматов обмена данными, и ускорить процесс интеграции форматов за счет использования единого формата представления знаний, что дает возможность автоматически строить схемы перевода из вновь зарегистрированного формата в любой из уже зарегистрированных в системе.

1 Предлагаемый подход и архитектура системы

Основной идеей предлагаемого подхода является введение «промежуточного» уровня – онтологии [1], хранящей знания о предметной области, т.н. нейтрального формата. Таким образом, знания, представленные в любом формате, будут храниться в специальной внутренней структуре, не зависящей ни от структуры формата, ни от платформы (см. рисунок 1). Подобная архитектура позволит осуществлять перевод из формата в формат естественным образом – как только построено соответствие между новым форматом и онтологией, обеспечивается возможность коммуникации между всеми уже зарегистрированными форматами.

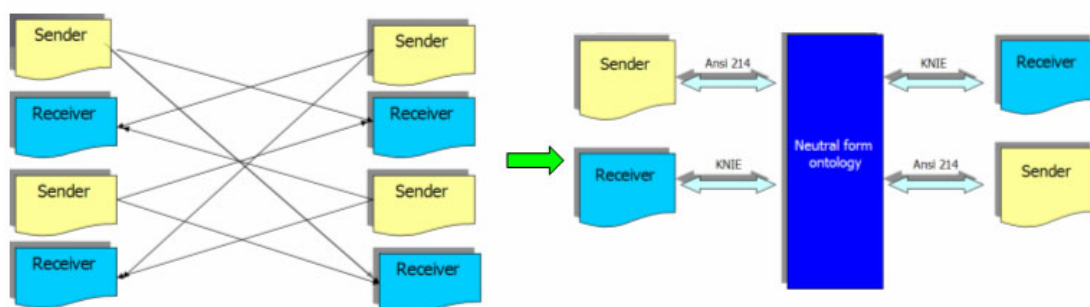


Рисунок 1 – Новая парадигма интеграции форматов обмена бизнес-данными

Таким образом, процесс работы системы выглядит следующим образом:

Каждый пользователь для регистрации должен представить ряд тестовых сообщений в новом формате. Также он может ввести ряд семантических правил и ограничений, описывающих его предметную область (например, правила, что он оперирует исключительно на территории Англии, или что его флот судов не превышает 100 штук, и т.п.) [2]. На основании онтологии предметной области, заданных правил, уже известных форматов обмена и тестовых сообщений, строится маппинг – схема соответствия между новым форматом и онтологией предметной области. В случае, если новый формат отличается от известных стандартных, запускается процедура автотмаппинга, когда на основе анализа сообщений пользователя строятся предположения о семантическом значении и связи с онтологией отличающихся полей, а также вычисляется степень достоверности каждого из подобных предложений.

Далее уточненный маппинг выдается эксперту для утверждения, и после правки и подтверждения новый формат считается зарегистрированным.

Теперь при посылке сообщений в этом формате они на основе выработанного маппинга преобразуются в нейтральный формат. И далее из нейтрального формата они могут быть преобразованы в любой зарегистрированный формат-приемник.

Также необходимо обеспечить механизм контроля и исправления ошибок, которые возникают внутри отдельных сообщений при транзакции (по оценке экспертов примерно 5-7 сообщений из ста содержат те или иные ошибки). Коррекция ошибочных сообщений осуществляется на основе данных из онтологии предметной области, онтологии используемого формата сообщения, семантических ограничений на поля сообщений, а также истории сообщений пользователя. Исправленное сообщение может либо автоматически быть перенаправлено для нового

прохода, либо, в случае, когда достоверность исправления невелика, или есть несколько равноправных вариантов, оно выдается эксперту для подтверждения.

Таким образом, данная архитектура системы позволяет автоматически регистрировать новые форматы, находить и исправлять отклонения в них, а также исправлять ошибки, встречающиеся в сообщениях при транзакциях, тем самым покрывая все основные элементы, требуемые для интеграции форматов (см. рисунок 2).

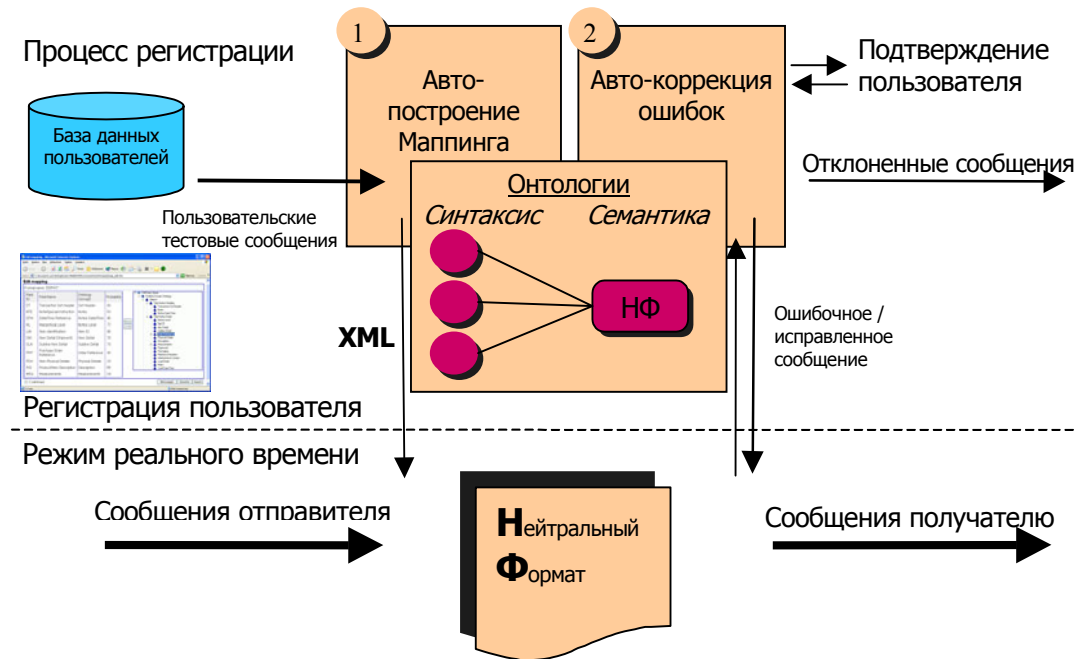


Рисунок 2 – Архитектура системы

2 Онтологии и маппинг

Центральным элементом системы является онтология нейтрального формата, содержащая знания о предметной области и интерпретирующая все основные форматы обмена данными. Разработанная система поддерживает спецификации документов, основанные на промышленных стандартах, таких как ANSI X12 (стандартный формат обмена бизнес-данными), EDIFACT (EDI ≈ Electronic data interchange), XML (множество правил для создания набора тэгов, описывающих бизнес-данные в структурированном формате) и ряд других.

В предлагаемом подходе используется два типа онтологий:

- Семантическая онтология (онтология нейтрального формата) – она существует в единственном виде на всю систему и содержит в себе знания в области логистики перевозок, покрывающие ряд форматов в пакетах стандартов ANSI X12 & EDIFACT (включая ANSI 214, ANSI 856, ANSI 810 и др.)
- Синтаксическая онтология – данная онтология требуется на каждый формат, и описывает общую структуру формата, а именно порядок полей и сегментов, их последовательность, возможность и условия образования циклов и т.п.

2.1 Семантическая онтология

Построенная в рамках системы семантическая онтология включает в себя несколько сот понятий, порядка 200 атрибутов (не считая перечислимые значения, общее число которых превышает 4 тысячи), и около 50 отношений. Данная онтология строилась экспертами в области логистики, и общие трудозатраты составили порядка 14 человеко-недель. В итоге онтология способна покрыть семантику ряда форматов обмена логистическими сообщениями, тем самым являясь хорошим базисом для интеграции.

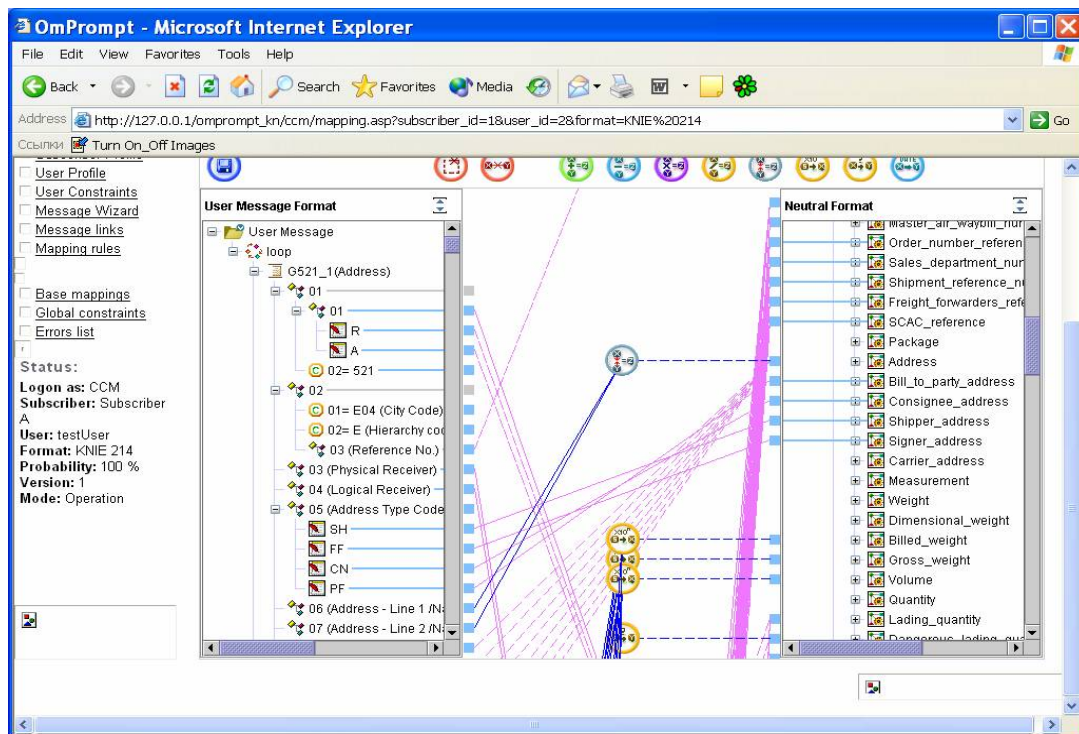


Рисунок 3 – Интерфейс для работы с синтаксической, семантической онтологиями и маппингом.

2.2 Синтаксическая онтология

Каждый стандарт определяет сегменты данных, состоящие из элементов данных, содержащих какую-то информацию (цена, номер модели продукта и т.п.). Последовательность из одного или более сегментов данных образует транзакцию, которая и передается по сети от одного торгового партнера к другому (обычно в зашифрованном виде).

Таким образом, синтаксическая онтология содержит в себе информацию о синтаксическом представлении формата, его структуре, включая список обязательных полей, последовательность полей, их повторяемость, разбиение на сегменты, правила, по которым поля в сообщении могут образовывать циклы.

Обычно синтаксическая онтология формата может быть построена в течение 2-3 дней.

2.3 Маппинг

Маппинг представляет собой множество правил, каждое из которых ставит в соответствие одно или несколько полей в исходном сообщении (форматная часть элемента – синтаксическая онтология) и предметной области («нейтральный формат» - семантическая онтология). Правила могут быть простые (поле A регистрируемого формата должно преобразовываться в поле X

нейтральной онтологии), условные (в зависимости от значения поле А, поле В может иметь разный смысл и маппиться на разные концепты в онтологии) и функциональные (некоторая функция применяется к набору полей из формата, и результат записывается в поле Х нейтрального формата. Примерами таких функций являются конкатенация, преобразования времени и даты, математические операции).

3 Автоматический перевод сообщений – автомаппинг

Признаком неполного соответствия формату является невыполнение одного из онтологических условий в построенной на основе текущего маппинга сцене, представляющей семантический смысл текущего сообщения [3, 4]. В этом случае система переходит в режим поиска отклонений в пользовательском формате.

Отклонения делятся на две группы – с добавлением семантики (то есть в сегменте появляются поля с семантикой, которой не имеет ни одно поле в общепринятом формате) и без добавления семантики.

На входе модуля находится множество из нескольких тестовых пользовательских сообщений в установленном формате. Предполагается, что все тестовые сообщения не содержат ошибок, а содержат только общие для всех сообщений отклонения от общепринятого стандарта.

На выходе модуля после обработки всех сообщений выдается маппинг пользовательского формата, то есть множество соответствий между полями пользовательских сообщений и понятиями предметной области.

При построении маппинга используется также данные из онтологии предметной области, данные из онтологии используемого формата сообщений, а также ограничения на поля, заданные пользователем во время регистрации.

Типы коррекций, которые поддерживает система, включают восстановление пропущенного значения поля на основе онтологии, определение неизвестного поля в сегменте, распознавание удаления поля, изменение семантики одного из полей, изменение порядка следования полей в сегменте и др.

Общий алгоритм нахождения маппинга состоит из следующих шагов –

- Нахождение места и типа отклонения на основе применения семантических правил и сравнения с онтологией синтаксиса формата
- Выработка допустимых вариантов коррекции на основе правил и тестовых сообщений пользователя
- Определение наилучшей коррекции из возможных, исходя из достоверности каждого из предложенных вариантов, а также комбинаторной вероятности в случае, когда для коррекции потребовалось сделать несколько последовательных допущений.

Результатом работы системы становится новый маппинг, где отмечены все найденные пользовательские отклонения от формата, а также указана достоверность для каждой из предложенных вариаций. В дальнейшем маппинг может редактироваться и утверждаться экспертом.

4 Коррекция ошибочных сообщений

На вход модуля поступает ошибочное сообщение, которое было определено в процессе исполнения транзакции путем применения семантических правил, описывающих допустимые варианты и ограничения формата. В случае несоответствия можно определить тип возникшей ошибки (или один из возможных типов), а также место возникновения ошибки.

В случае, если возможно автоматическое исправление ошибки, на выходе модуля появляется исправленное сообщение. В противном случае, модуль направляет сообщение эксперту, в котором указывается исходное сообщение, тип и место возникновения ошибки, а также, список

возможных вариантов исправления ошибки (в случае, когда такие варианты есть, но ни одному из вариантов нельзя отдать предпочтение).

Типы ошибок, которые способна исправлять система, включают пропущенное значение поля, некорректная длина значения поля (например, для даты или времени), недопустимое значение из списка, выход за допустимый диапазон, значение, не соответствующее набору шаблонов (например, для почтового индекса) и др.

Ошибочный элемент сцены выявляется при построении сцены на основе маппинга. Ошибочный элемент может быть либо атрибутом (если ошибка возникла в поле, соответствующем нетриггерному элементу маппинга), либо объектом (в поле триггерного элемента – т.е. элемента, который влияет на смысл интерпретации другого элемента формата).

В последнем случае множество предполагаемых замен ограничено «братьями» данного объекта в онтологии. В случае ошибки в атрибуте, для коррекции применяется ряд эвристических методов изменения значения:

- замена значения на одно из значений, упоминавшихся в предыдущих сообщениях
- изменение способа кодирования (для атрибутов даты, числа)
- символьные алгоритмы замены неверного значения на похожее (для текстовых атрибутов) на основе алгоритмов нечеткого поиска

Сам же процесс коррекции значения состоит из следующих этапов:

- 1) Поиск возможных вариантов для замены.
- 2) Отбрасывание вариантов, которые не удовлетворяют онтологическим ограничениям.
- 3) Определение ценности каждого варианта и выбор наиболее вероятного.

В итоге исправленный вариант должен удовлетворять всем онтологическим ограничениям. Величина коррекции должна быть минимальной и в то же время максимально соответствовать тестовым сообщениям пользователя [5].

Если степень достоверности исправленного варианта достаточно велика, то коррекция осуществляется автоматически, в противном случае несколько наилучших вариантов, из которых система затрудняется сделать выбор, представляются эксперту.

5 Основные результаты

Согласно предложенному подходу была разработана пилотная версия системы, обладающая следующими характеристиками:



Рисунок 4 – Процесс преобразования форматов

- Поддерживаемые форматы обмена бизнес-данными: 12 форматов, включая ANSI 810, ANSI 856, ANSI 214, EDIFACT IFTSTA, EDIFACT INVOIC, EDIFACT DESADV, XML Voice, XML Fax, а также ряд специфичных клиентских форматов.

- Объемы онтологии: ~ 350 объектов и отношений, более 4000 значений перечислимых атрибутов.
- Скорость обработки сообщений: 100 сообщений на основе правил маппинга преобразуются за 13-18 сек.
- Скорость нахождения допустимых исправлений для ошибочного сообщения: ~ 1-3 сек.
- Точность преобразования в среднем составляет 88% для перевода разных форматов, 96% для различных вариаций одного формата.
- В качестве механизма представления маппинга используется язык преобразований XSLT, XML Schema используется для хранения правил и семантических ограничений, накладываемых на формат пользователя. Подобная архитектура позволяет поддерживать основные промышленные стандарты интеграции приложений в масштабах предприятий и интеграции процессов электронного бизнеса, в частности IBM MQSeries.

Заключение

К основным результатам проекта можно отнести следующие возможности, обеспечиваемые системой:

- Предложен новый подход к интеграции различных стандартов обмена бизнес-сообщениями, позволяющий существенно ускорить и упростить процесс установления взаимосвязей между партнерскими компаниями.
- Предложен механизм для автоматического определения и исправления возможных отклонений в пользовательских форматах.
- Достигнута высокая эффективность в автоматическом нахождении и коррекции ошибок в сообщениях известных форматов.
- Обеспечена возможность пополнения онтологии и подключения новых форматов с помощью экспертов предметных областей, без необходимости дополнительного программирования.

Рассматриваемая система в дальнейшем станет основой для автоматической интеграционной платформы, помогающей автоматизировать и налаживать совместный бизнес компаний-партнеров.

Список литературы

- [1] Batishchev S.V., Ivkushkin C.V., Minakov I.A., Rzevski G.A., Skobelev P.O. MagentA Multi-Agent Systems: Engines, Ontologies and Applications // Proc. of the 3rd Intern. Workshop on Computer Science and Information Technologies CSIT'2001, Ufa, Russia, 21-26 September, 2001. – Ufa State Aviation Technical University – Institute JurInfoR-MSU, Vol. 1: Regular Papers, 2001, pp. 73-80.
- [2] V.Andrejev, S.Batishchev, K.Ivkushkin, I.Minakov, G.Rzevski, A.Safronov, P.Skobelev. MagentA Multi-Agent Engines for Decision Making Support // International Conference on Advanced Infrastructure for Electronic Business, Science, Education and Medicine on the Internet (ISBN 88-85280-63-3), 29 July – 4 August 2002, L'Aquila, Italy, pp. 64-76.
- [3] Андреев В.В., Ивкушкин К.В., Карягин Д.В., Минаков И.А., Ржевский Г.А., Скобелев П.О., Томин М.С. Разработка мультиагентной системы понимания текста. // Тр. III Международной конф. «Проблемы управления и моделирования в сложных системах». – Самара: СНЦ РАН, 2001. С. 489-495.
- [4] В. Андреев, С. Вольман, К. Ивкушкин, Д. Карягин, И. Минаков, А. Пименов, П. Скобелев, М. Томин. Разработка мультиагентной системы интеллектуальной обработки и классификации документов// Труды V Международной конференции по проблемам управления и моделирования сложных систем (Самара, 17-21 июня 2003). – Самара: СНЦ РАН, 2003. С. 317-323.
- [5] Андреев В.В., Волхонцев Д.В., Ивкушкин К.В., Карягин Д.В., Минаков И.А., Ржевский Г.А., Скобелев П.О. Мультиагентная система извлечения знаний. // Тр. III Международной конф. по проблемам управления и моделирования сложных систем (Самара, 4-9 сентября 2001). – Самара: СНЦ РАН, 2001. С. 206–212.

РАЗВИТИЕ ЭЛЕМЕНТОВ САМООРГАНИЗАЦИИ И ЭВОЛЮЦИИ В МУЛЬТИАГЕНТНОМ ПОРТАЛЕ СОЦИОКУЛЬТУРНЫХ РЕСУРСОВ САМАРСКОЙ ОБЛАСТИ

В.В. Андреев¹, О.И. Лахин³, И.А. Минаков², А.Н. Сальков³, П.О. Скобелев²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²ООО НПК «Маджента Девелопмент»
443110, Самара, ул. Осипенко, 1а

contact@magenta-technology.ru

тел: +7 8462 706684, +7 8462 706685 факс: +7 8462 706685

³ООО НПК «Генезис Знаний»

443011, Самара, ул. Ново-Садовая, д. 221, офис 206

salkov@kg.ru

тел/факс: +7 (8462) 34-87-57, 17-30-69

Ключевые слова: открытые мультиагентные системы, самоорганизация и эволюция, извлечение знаний, понимание текста, социокультурные ресурсы

Abstract

The paper discusses new approach for self-organization and evolution applied to Internet-portal of social and cultural resources of the Samara region. This approach develops and advances the ideas of the Semantic Web, basing on the multi-agent technologies, knowledge management and text understanding algorithms.

Введение

Быстрое развитие сети Интернет сделало для пользователя доступным огромное количество информационных ресурсов, а также Интернет – порталов, предоставляющих целый комплекс информационных услуг. Но гигантские объемы имеющейся информации становятся причиной ряда проблем, в первую очередь связанных с необходимостью быстро и корректно обрабатывать и получать релевантные данные и знания.

К числу основных проблем относится задача индивидуализации поиска, когда каждый пользователь имеет возможность указать, какие именно предметные области/темы для него являются наиболее важными, и получить только актуальную информацию. Для решения этой задачи необходимо наличие интеллектуальных средств, способных понять потребности и желания каждого отдельного пользователя, а также механизмы работы с информационными ресурсами, позволяющие оперативно отреагировать на его запросы.

Второй серьезной проблемой является «пассивность» информационных ресурсов – когда имеющиеся данные никак не связаны между собой, и появляющиеся различные статьи и сайты на одну и ту же тему, зачастую взаимно дополняющие друг друга, существуют независимо, без коммуникации друг с другом (иногда даже не зная о существовании друг друга), не имея возможности организоваться в некое «сообщество по интересам», а также сами про-активно искать пользователей, проявляющих интерес к данной предметной области.

Поэтому для решения указанных проблем W3 Consortium¹ разрабатывает концепцию Семантической Сети. Семантическая Сеть (Semantic Web) – расширение современной сети, в которой каждому информационному ресурсу присваивается формализованный смысловой описатель, представляющий семантическое значение этого ресурса. Суть этой идеи состоит в том, чтобы научить, наконец-то, компьютеры “понимать” семантическое содержание Интернет-сайтов, то есть распознавать и интерпретировать смысл (значение) той информации, которая находится на странице [1].

Для решения данной проблемы нами предлагается подход работы со знаниями, основанный на принципах самоорганизации и эволюции, и реализуемый с помощью мультиагентной технологии. В настоящей статье мы рассмотрим особенности его применения на примере создание мультиагентного портала социокультурных ресурсов.

1 Мультиагентный портал социокультурных ресурсов Самарской области

Для интеграции деятельности общественных организаций и объединений, а также стимулирования процессов самоорганизации людей в таких сообществах был создан Интернет – портал социокультурных ресурсов Самарской области [2].

Портал представляет собой открытую и развивающуюся систему, позволяющую пользователям всех категорий не только получить интересную для них информацию, но и принять активное участие в повседневной жизни и развитии портала. Для обеспечения указанной функциональности каждому пользователю и ресурсу ставятся в соответствие агенты - программные объекты, постоянно следящие за изменениями системы портала и стремящиеся найти наиболее интересные ресурсы на основе динамически изменяющейся карты интересов и знаниях о пользователе. При этом знания представлены онтологией портала, в которой хранятся социокультурные ресурсы – кванты и связи культурной сферы деятельности между ними. Каждый пользователь может не только добавить новый квант, но и развивать существующую онтологию, путем построения семантических дескрипторов (описаний) квантов.

Агенты в дальнейшем используют онтологий социокультурных ресурсов для поиска наиболее интересных для пользователя квантов, при этом, накапливая информацию о действиях пользователя и заботясь о том, чтобы картина его исходных предпочтений постоянно уточнялась и корректировалась с каждым его шагом в системе.

Таким образом обеспечивается интеллектуальное взаимодействие с пользователем, динамическое формирование страниц портала с учетом интересов пользователя, индивидуализация работы портала под каждого пользователя и открытость системы.

На сегодняшний день в портале социокультурных ресурсов Самарской области насчитывается более 450 квантов, и примерно в полтора раза больше – связей между ними [3]. Наличие такого объема связанной информации, а также ее постоянный рост порождает следующие проблемы:

- усложняется процесс реорганизации онтологии портала, то есть динамического изменения связей между квантами;
- затрудняется возможность адекватной реакции на изменение поведения пользователей;
- усложняется доступ к актуальной для пользователя информации.

Для решения указанных проблем необходимо применение механизмов самоорганизации и эволюции в системе портала.

¹ Международный индустриальный консорциум, целью которого является разработка общих стандартов сети Интернет.

2 Механизм самоорганизации в портале социокультурных ресурсов Самарской области

Для того чтобы система портала адекватно и своевременно реагировала на изменение интересов пользователей, необходимо наличие механизма самоорганизации, то есть наличие механизма автономной реорганизации связей в системе. Самоорганизация в портале возможно благодаря тому, что агенты пользователей и ресурсов способны самостоятельно воспринимать ситуацию и принимать решения об изменении связей между квантами.

В качестве механизма, обеспечивающего процессы самоорганизации, используется метод компенсаций, когда каждый информационный ресурс имеет некий уровень энергии (виртуальные деньги), отражающие его информационную ценность\популярность. В процессе переговоров информационные ресурсы соревнуются за возможность быть показанным тому или иному пользователю, оплачивая установление связи, а пользователь, в свою очередь, реагируя на интересность представленной ему информации, дает деньги наиболее заинтересовавшим его квантам, посредством выбора той или иной статьи.

Таким образом, задача агента кванта – распределить свои деньги так, чтобы получить максимальный доход. В случае траты всех денег агент погибает, а соответствующий квант удаляется из системы. Так, например, рискованный агент может инвестировать все свои деньги в связь с самым посещаемым квантом, но за счет высокой оплаты связи разориться и отказаться от связи или погибнуть в системе. Однако он может вовремя понять ситуацию и поменять свою стратегию на менее рискованную.

Более того, агенты могут организовывать связи на основе подобия содержимого, то есть в результате матчнга (сравнительной схожести) их семантических дескрипторов, таким образом, образуя группы по определенным темам.

3 Механизм эволюции в портале социокультурных ресурсов

Для того чтобы обеспечить пользователей новыми и интересными ресурсами и чтобы портал мог адекватно и своевременно реагировать на изменение поведения пользователей, например при снижении уровня посещаемости портала, необходимо наличие механизма эволюции и развития.

При этом эволюция представляется, как неограниченная последовательность процессов самоорганизации в условиях постоянно меняющихся внешних условий, в результате которой из простой "плоской" сети квантов портала образуется более сложная комбинация групп (кластеров) квантов, а также появляются совершенно новые кванты (агент рекомендует автору изменить содержимое заметки и появляется статья с новым содержанием, агент находит информацию в сети, которую просит подключить к своему тексту и т.д.).

Агент ресурса, у которого понижается уровень посещаемости, может:

- инициировать своего хозяина, который его создал, чтобы он поменял семантический дескриптор, при этом агент может предоставить список рекомендуемых изменений на основе анализа интересов и смысла своего ресурса
- самостоятельно по своему семантическому дескриптору установить связь с другими схожими ресурсами, которые находятся за пределами системы портала;
- пригласить людей посетить портал.

В случае изменения поведения пользователей, например, при появлении нового кванта, уровень посещаемости которого резко возрастает, агент может автоматически изменить стратегию, пересмотреть свои связи и по возможности установить связь с этим квантом.

Таким образом, данные механизмы позволяют оперативно реагировать на любые изменения в структуре портала, будь то появление новой информации или смена интересов групп пользователей, давая в любом случае наиболее адекватную информацию, динамически подстраиваясь под текущие интересы посетителей портала.

4 Применение систем извлечения знаний и понимания текста в задачах самоорганизации и эволюции

Для поддержки механизмов самоорганизации и эволюции используются системы извлечения знаний [3] и понимания текста [4].

Система понимания текста применяется для формирования семантических дескрипторов каждого информационного ресурса, давая возможность представить семантический смысл каждого кванта портала. Также на ее основе реализуются механизмы поиска похожих квантов, позволяя образовывать группы схожих материалов, обеспечивая самоорганизацию информации в портале.

Механизмы эволюции проявляются в возможности системы понимания текста пополнять онтологию новыми понятиями, знаниями, которые встречаются в поступающих информационных материалах, тем самым увеличивая знания агентов о том мире, в котором они существуют, и давая возможность использовать вновь обретенные знания в логике рассуждений. (Например, если на сайт добавляется новый квант, рассказывающий о предстоящей выставке, а само понятие «выставка» не представлено в онтологии, то на основе понимаемого контекста будет сделан вывод, что выставка – вид социокультурного события, и в дальнейшем этот факт, добавленный в онтологию, будет использоваться при анализе и выработке предложений пользователям).

Для поддержки механизма самоорганизации применяется система извлечения знаний, позволяющая адекватно реагировать на поведение пользователей, динамически уточняя карту их интересов на основе наблюдения за действиями пользователя.

Так, агенты квантов и пользователей портала могут образовывать связанные группы - кластеры, при этом пользователи могут объединяться на основе общих интересов и увлечений, а кванты на основе подобия содержимого их семантических дескрипторов. Кроме того, кванты могут образовывать группы наиболее посещаемых ресурсов по определенной тематике. (Например, группа информационных материалов, посвященных событиям конца 19-го века, могла быть задана вручную в процессе построения онтологии, а могла и образоваться самостоятельно на основе механизма кластеризации, когда стало понятно, что многие материалы имеют эту общую особенность, а данная тема востребована посетителями портала).

Также механизм извлечения знаний применяется для обеспечения эволюционного развития агентов, когда посредством анализа удачных и неудачных стратегий продвижения информационных квантов пользователям, агенты динамически меняют стратегии своего поведения, чтобы лучше быть приспособленным к имеющейся ситуации и быть точнее ориентированными на те или иные группы пользователей.

Системы извлечения знаний и понимания текста обеспечивают основные механизмы работы со знанием – представление, извлечение, поиск, анализ, пополнение. На основе этих механизмов и осуществляются процессы самоорганизации и эволюции, рассмотренные нами на примере социокультурного портала.

5 Идеологические аспекты механизмов самоорганизации и эволюции

Рассматриваемые механизмы самоорганизации и эволюции на теоретическом уровне подробно разобраны в работах Г. Хакена [5] и И. Пригожина [6].

В указанных работах изучаются процессы самоорганизации и эволюции в сложных открытых системах, к которым относится также и портал социокультурных ресурсов, так как структура портала может динамически изменяться, а число элементов достаточно высоко.

Под самоорганизацией понимается спонтанное образование устойчивых организаций, структур при определенных условиях, например, образование групп агентов по определенной тематике при наличии достаточного количества квантов со сходными семантическими дескрипторами. При этом эволюция представляется, как неограниченная последовательность про-

цессов самоорганизации в условиях постоянно меняющихся внешних условий. Для портала социокультурных ресурсов такими условиями, например, являются динамика появления в портале новых ресурсов, уровень посещаемости портала, типы и интересы пользователей.

Кроме того, в сложных открытых системах наблюдаются различные феномены, которые мы надеемся наблюдать в портале:

- осцилляция связей между квантами,
- катастрофы - лавинообразный пересмотр связей, например, при появлении нового интересного кванта в портале,
- аттракторы – устойчивые структурообразования, например группы квантов по определенной тематике.

Заключение

В настоящей статье рассмотрены основные аспекты построения кардинально нового инструмента с использованием развитых элементов самоорганизации и эволюции, позволяющие решить такие проблемы, как поиск актуальной для пользователя информации и обеспечение «про-активности» информационных ресурсов.

Использование Интернет-порталов «по интересам» дает возможности индивидуализации поиска и работы со знаниями для каждого пользователя на основе его интересов и предпочтений, определяемых автоматически и динамически подстраивающихся. Алгоритмы извлечения знаний и понимания текста позволят интегрировать различные информационные ресурсы, создавая сообщества «по интересам» и связывая информацию из различных информационных источников, тем самым продвигая идеи Semantic Web в новом направлении, где знание автоматически определяется на основе текста сайта, автоматически коммуницирует и преобразует себя на основе взаимодействия с пользователем, и в итоге создает среду, в которой каждый пользователь оперативно может получить всю необходимую информацию в удобной форме с множеством гиперсвязей на интересующие его материалы.

Список литературы

- [1] Батищев С.В., Искварина Т.В., Скобелев П.О. Методы и средства построения онтологий для интеллектуализации сети Интернет // Известия Самар. науч. центра РАН. – 2002. – Янв. – июнь. – С. 126-137.
- [2] Батищев С.В., Генералова Г.Д., Горбунова Т.Ф., Куруленко Э.Я., Лахин О.И., Скобелев П.О. Мультиагентный Интернет-портал по социокультурным ресурсам Самарской области // Труды 4-й Международ. конф. по пробл. упр. и моделир. сложных систем, Самара, 17-24 июня 2002. – Самара: СНЦ РАН, 2002. С. 449-458.
- [3] Андреев В.В., Волхонцев Д.В., Ивкушкин К.В., Карягин Д.В., Минаков И.А., Ржевский Г.А., Скобелев П.О. Мультиагентная система извлечения знаний. // Труды 3-й Международной конференции по проблемам управления и моделирования сложных систем.
- [4] Андреев В.В., Ивкушкин К.В., Карягин Д.В., Минаков И.А., Ржевский Г.А., Скобелев П.О., Томин М.С. Разработка мультиагентной системы понимания текста. // Труды 3-й Международной конференции по проблемам управления и моделирования сложных систем (в печати).
- [5] Хакен Г. Синергетика: Иерархия неустойчивостей в самоорганизующихся системах и устройствах: Пер. с англ. – М.: Мир, 1985. - 423 с., ил. – С. 86
- [6] Николис Г., Пригожин И. Самоорганизация в неравновесных системах. М., 1979. – 289 с.

МУЛЬТИАГЕНТНАЯ АРХИТЕКТУРА ОБСЛУЖИВАНИЯ ЗАПРОСОВ К БАЗАМ ДАННЫХ УДАЛЕННОГО ВОСПРИЯТИЯ

В.В. Девятков

Московский государственный технический университет им. Н.Э. Баумана
105005, Москва, ул. 2-ая Бауманская, 5
deviatkov@iu3.bmstu.ru
тел: +7 (095) 263-66-25

Ключевые слова: интеллектуальный агент, нечеткое исчисление, синтаксическое распознавание, интеллектуальные датчики, удаленное восприятие, база данных удаленного восприятия

Abstract

Sensor networks are being widely deployed for measurement, detection and surveillance applications. Sensor network is potentially global network of smart sensing nodes, with web cams or other monitoring devices, and organizing nodes. In these new applications, users need to query very large collections of devices in ad-hoc access. Most existing systems rely on a centralized system for collecting device data. These systems lack flexibility because data is extracted in a pre-defined way; also they do not scale to a large number of devices because large volumes of raw data are transferred. In our concept of remote sensing database, distributed query execution techniques are applied to leverage the computing and capabilities of devices and to reduce communication. Sensor data represented as time 2D or 3D series. The query system depends on ability to recognize property of series. It is known, that recognition can be considered as the syntactic analysis in some language. It is known also, that languages in which fuzzy sets lay in basis of semantics, can serve as the convenient tool for recognition of fuzzy and irregular series. The languages based on fuzzy rules are popular. Present paper represents the approach to syntactic recognition of series, based on the fuzzy situational calculus used by hierarchically organized community of intellectual agents. Agents may be distributed. Each agent deals with the volume of the knowledge submitted in fuzzy situational calculus as his ontology. All ontologies make up distributed remote sensing database that provide the means to query present and historical sensor-based data. In paper within the framework of the accepted concept of fuzzy hierarchical multi-agent syntactic recognition the essence of original querying is considered. Illustrative examples are resulted.

Введение

Прогнозы использования для анализа ситуаций в самых различных физических средах беспроводных сетей сбора и обработки распределенной информации с удаленных датчиков чрезвычайно восторженные [1]. В развитых странах в последние годы ведутся интенсивные исследования в этом направлении [2]. Беспроводные датчики становятся все дешевле, что позволяет расставлять их в самых экзотических местах и в большом количестве. Число датчиков может достигать нескольких тысяч [3]. Датчики становятся все разнообразнее и интеллектуальнее. Сбор и обработка распределенной информации с удаленных датчиков, может быть централизованной, децентрализованной или комбинированной. При большом числе датчиков централизованная обработка становится слишком трудоемкой. Предпочтительными становятся децентрализованные и комбинированные сбор и обработка информации. Области применения беспроводных или комбинированных сетей расширяются от достаточно традиционных типа анализа освещенности, температуры, движения, контроля доступа до весьма экзотических типа контроля состояния мобильного телефона, автомобиля, личного кошелька с деньгами, наблюдения за детьми на детской площадке или за собаками, гадящими в неполюженном месте. Во

всех случаях информация поступает в базу данных, которую будем называть базой данных удаленного восприятия (БДУВ). В соответствии с типом сбора и обработки информации с датчиков БДУВ может быть централизованной, децентрализованной или комбинированной. В предлагаемой статье БДУВ распределена между агентами, собирающими и обрабатывающими информацию. Агенты могут находиться в различных узлах сети и общаются друг с другом с целью сбора, передачи и обработки информации. Они также обслуживают различных потребителей, которые обращаются к БДУВ с запросами с целью получения той или иной информации о среде, где расположены датчики. Часть БДУВ, принадлежащая конкретному агенту, является онтологией [4]. Основное внимание в статье уделяется принципам формирования и использованию онтологий сообществом интеллектуальных агентов для ответа на запросы пользователей к БДУВ. В настоящей статье под онтологией понимается формализованное описание каких-либо свойств образов (сигналов, изображений, сцен). Под ситуацией понимается некоторое выводимое (устанавливаемое) в процессе мультиагентного распознавания свойство образов, которое они приобретают в определенное время или в определенном месте. Та или иная ситуация может определяться специальными параметрами (переменными). В качестве языка описания онтологий используется нечеткое ситуационное исчисление, изложенное в [5]. Ответ на запрос рассматривается как вывод в этом исчислении.

1 Синтаксические принципы формирования свойств образов

Нечеткая распределенная между агентами БДУВ состоит из двух частей – описания текущих параметров образов и описания свойств образов. Обе эти части используются для анализа ситуаций. Текущие параметры регулярно обновляются. Рассмотрим некоторые принципы формирования онтологий характеризующих поведение образов объектов во времени и пространстве. Образом называется некоторая физическая величина (предмет, явление), имеющая хотя бы один параметр, который содержит информацию и называется информативным. Эта информация может быть определена (измерена) для каждого параметра как его значение из определенного множества значений. Значения информативных параметров образов используются для суждений о свойствах физической величины (явлении), например, с целью диагностики заболеваний сердца, если образ является кардиограммой, или для целей управления скоростью вращения двигателей насосов, если образы представляют графики изменения давления жидкости в трубопроводах. Параметры образов могут быть непрерывными и дискретными. Непрерывный параметр задан на всей области определений X . Дискретный параметр задан только в отдельных точках области определения, например, через один и тот же шаг дискретизации Δ_x , начиная с некоторого начального значения: $x_0 = x$, $x_1 = x + \Delta_x$, $x_2 = x + 2\Delta_x$, $x_3 = x + 3\Delta_x, \dots$.

Шаги дискретизации могут быть различными. Цифровые образы, как правило, имеют только дискретные параметры. Для простоты рассмотрим случай, когда образ имеет только два параметра $x \in X$ и $y \in Y$. Параметры находятся в некотором отношении $R(x, y) \in X \times Y$.

На рисунке 1 представлен пример цифрового образа (сигнала), имеющего два параметра x и y . Шаг дискретизации параметра x равен Δ_x , параметра y - Δ_y . Поставим в соответствие каждому отношению $R(x_i, y_j)$ вершину графа, помеченную символом N_i , $i = 0, 1, \dots, 12$. (вершины будем обозначать кружком). Самую левую вершину пометим также символом $S_3 = N_0$. Каждую пару соседних вершин $i, i+1, i=0, 1, 2, \dots, 11$ соединим дугой, направленной от вершины i к вершине $i+1$. Дугу, направленную от вершины i к вершине $i+1$ пометим символом a_i в алфавите $A_3 = \{a_0, a_1, a_2, a_3, \dots, a_m\}$. В результате получим граф сегментации цифрового сигнала, показанный на рисунке 2. Если выписать обозначения всех дуг слева направо, то получим последовательность символов $a_0a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}$. Дуги, имеющие одно и то же направление и длину, переобозначим одним и тем же символом, встречающимся первый раз в последовательности $a_0a_1a_2a_3a_4a_5a_6a_7a_8a_9a_{10}a_{11}$ при просмотре ее слева направо. В результате получим последовательность $a_0a_1a_2a_3a_4a_3a_6a_7a_8a_9a_6a_0$ и синтаксический граф сигнала, показанный на рисунке 3.

Эта последовательность может рассматриваться как слово или предложение некоторого языка $L_3 = L_3(G_3)$, порождаемого автоматной грамматикой $G_3 = \{V_3, A_3, P_3, S_3\}$, $V_3 = \{N_1, N_2, N_3, N_4, N_5, N_6, N_7, N_8, N_9, N_{10}, N_{11}\}$, $A_3 = \{a_0, a_1, a_2, a_3, a_4, a_6, a_7, a_8, a_9\}$, $P_3 = \{S_3 \rightarrow a_0 N_1, N_1 \rightarrow a_1 N_2, N_2 \rightarrow a_2 N_3, N_3 \rightarrow a_3 N_4, N_4 \rightarrow a_4 N_5, N_5 \rightarrow a_3 N_6, N_6 \rightarrow a_6 N_7, N_7 \rightarrow a_7 N_8, N_8 \rightarrow a_8 N_9, N_9 \rightarrow a_9 N_{10}, N_{10} \rightarrow a_6 N_{11}, N_{11} \rightarrow a_0\}$.

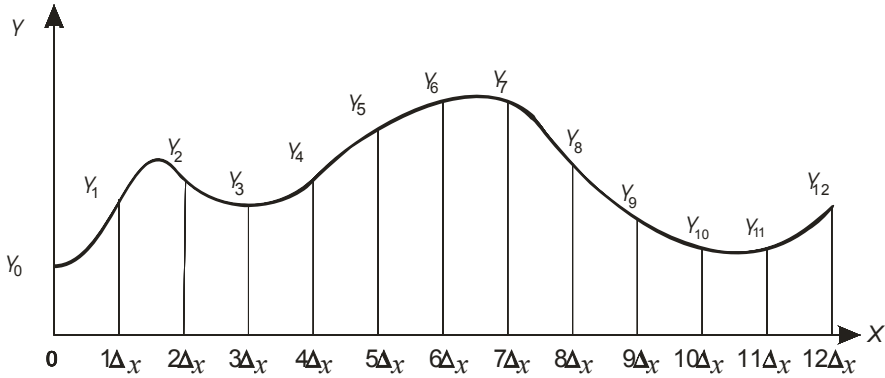


Рисунок 1 – Цифровой сигнал. Δ_x - шаг дискретизации по оси абсцисс.

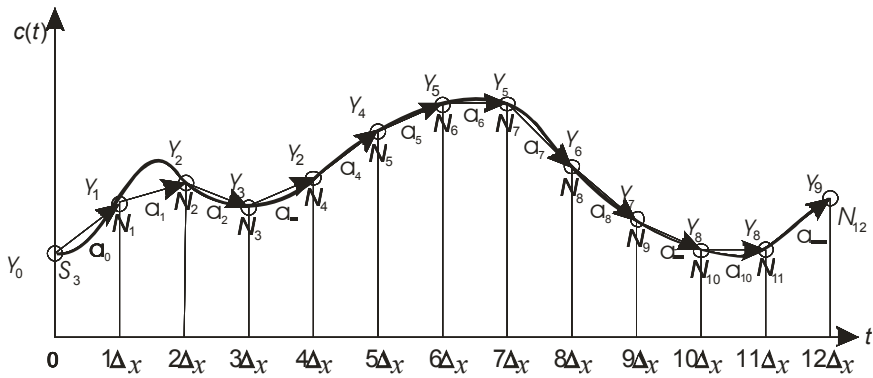


Рисунок 2 – Сегментация цифрового сигнала.

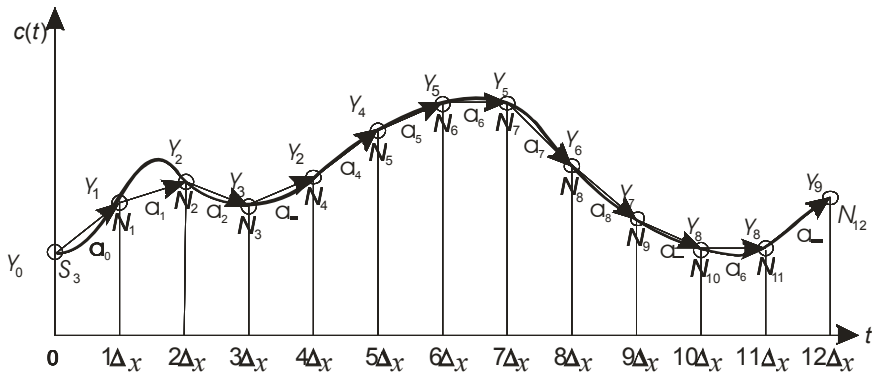


Рисунок 3 – Синтаксический граф цифрового сигнала.

Для того, чтобы перейти к представлению текущих и общих свойств образов, построим по четкой грамматике G_3 нечеткую грамматику, базируясь на следующих принципах.

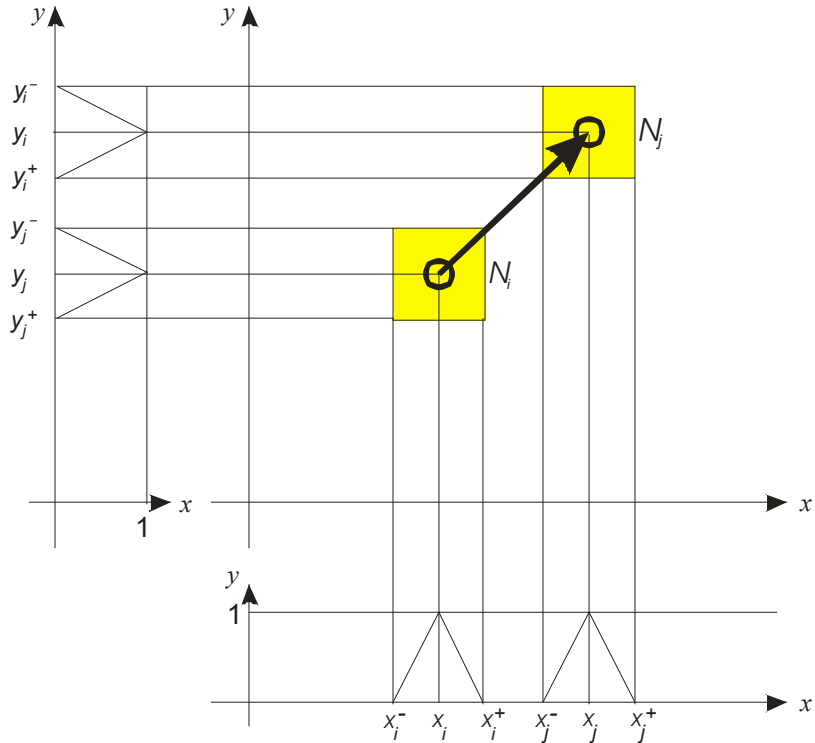


Рисунок 4 – Функции принадлежности вершин синтаксического графа.

Каждой дуге синтаксического графа соответствуют две инцидентные вершины N_i и N_j (рисунок 4). Координатами вершины N_i являются x_i и y_i , а координатами вершины N_j - координаты x_j и y_j . Допустим, что координаты обеих вершин могут изменяться в пределах прямоугольных областей, например, квадратов, приведенных на рисунке 4 с шагами дискретизации соответственно δ_x и δ_y . Такое допущение означает, что вместо одной вершины N_i с координатами x_i, y_i будем иметь множество вершин $N_{ni} \in B(N_i)$ с координатами, изменяющимися в пределах области, очерчивающей вершину N_i (на рисунке 4 это квадрат), причем

$$n = \frac{(x_i^- - x_i^+)}{\delta_x} \cdot \frac{(y_i^- - y_i^+)}{\delta_y}.$$

Вместо одной вершины N_j с координатами x_j, y_j при тех же диапазонах отклонения значений параметров и тех же значениях δ_x и δ_y будем иметь множество вершин $N_{nj} \in B(N_j)$ с координатами, изменяющимися в пределах квадрата, очерченного вокруг вершины N_j , а вместо одной дуги a_i , ведущей из вершины N_i в вершину N_j - множество всех дуг $a_{li} \in A(a_i) = \{(N_{ri}, N_{sj}) \mid N_{ri} \in B(N_i), N_{sj} \in B(N_j), l \leq n^2, r \leq n, s \leq n\}$, соединяющих каждую вершину множества $B(N_i)$ с каждой вершиной множества $B(N_j)$

Проекциями вертикальных сторон квадратов, очерчивающих множества $B(N_i)$ и $B(N_j)$, на ось X будут соответственно x_i^- и x_i^+ , x_j^- и x_j^+ , а на ось Y - y_i^- и y_i^+ , y_j^- и y_j^+ . Зададим четыре треугольные функции принадлежности $\mu_i(x), \mu_i(y), \mu_j(x), \mu_j(y)$, которые заданы на универсумах

X и Y , тройками значений соответственно в точках $\{x_i^-, x_i, x_i^+\}$, $\{x_j^-, x_j, x_j^+\}$, $\{y_i^-, y_i, y_i^+\}$, $\{y_j^-, y_j, y_j^+\}$ (см. рисунок 4). Каждая из этих функции определяется следующим выражением.

$$\mu(u) = \left\{ \begin{array}{l} 1, \text{ если } u = u_k \\ 0, \text{ если } u = u_k^- \\ 0, \text{ если } u = u_k^+ \\ \frac{u}{u_k - u_k^-} - \frac{u_k^-}{u_k - u_k^-}, \text{ если } u_k^- < u < u_k \\ 0, \text{ если } x = x_k^+ \\ \frac{-u}{u_k^+ - u_k} - \frac{u_i^+}{u_k^+ - u_k}, \text{ если } u_k < u < u_k^+ \end{array} \right\}.$$

Здесь $u = x, y, k = i, j$. Эти функции определяют меру близости координат вершин графа к «наилучшим координатам», которым соответствует значение функций принадлежности, равное 1.

Будем полагать, что функция принадлежности каждой дуги $a_{li} \in A(a_i)$, инцидентной вершинам $N_{ni} \in B(N_i)$ и $N_{nj} \in B(N_j)$, определяется следующим образом:

$$\mu_{A(a_i)}(a_{li}) = \min \{ \mu_i(x_{ni}), \mu_i(y_{ni}), \mu_j(x_{nj}), \mu_j(y_{nj}) \}.$$

Нечеткая грамматика $G_F = \{V, A_F, P_F, S_F\}$ получается из четкой грамматики $G = \{V, A, P, S\}$ следующим образом:

$$A_F = \bigcup_i A(a_i).$$

Единственный начальный нетерминальный символ четкой грамматики заменяется множеством начальных нетерминальных символов

$$S_F = B(b_0).$$

Множество правил P_F нечеткой грамматики G_F определяется выражением

$$P_F = \{ a_{li}(N_i \rightarrow a_{li}N_{i+1}), \mu(N_i \rightarrow a_{li}N_{i+1}) = \mu_{A(a_i)}(a_{li}), a_{li} \in A(a_i), i=0,1,\dots,m, l \leq n^2 \}.$$

Нечеткая автоматная грамматика G_F порождает нечеткий язык $\{L(G_F), R_{L(G_F)}\}$

$$L(G_F) = \{ l^* \mid l^* = a_{l_0}a_{l_1}a_{l_2}\dots a_{l_m}, a_{li} \in A(a_i), l \leq n^2 \},$$

$$R_{L(G_F)} = \{ \mu_{L(G_F)}(l^*) \mid l^* = a_{l_0}a_{l_1}a_{l_2}\dots a_{l_m}, a_{li} \in A(a_i), \mu_{L(G_F)}(l^*) = \min_{a=a_{l_0}, a_{l_1}, \dots, a_{l_m}} \{ \mu_{A(a_i)}(a_{li}) \}$$

2 Онтология свойств образов

Известно [6], что правила любой, в том числе нечеткой, грамматики G_F легко представить правилами $p(G_F)$ адекватного исчисления, в частности, нечеткого. Первоначальное построение грамматики является удобной формой хорошо формализуемого способа формирования в нечетком исчислении последовательностных (т.е. наблюдаемых во времени и/или пространстве) свойств образов. Онтология свойств образов состоит из описания последовательностных свойств образов $p(G_F)$, мгновенных свойств $\varphi(B(N_j))$, задаваемых на каждом множестве $B(N_j)$ и

фактических значений измеряемых параметров (универсумов значений) U . Онтология хранится в БДУВ в определенных форматах и распределена между агентами. Каждый из агентов решает некоторую частную задачу в соответствии со своей ролью. Вместе они решают общую задачу ответа на запрос. Каждый агент имеет входной и выходной порты, базу знаний (онтологию) и решатель или машину вывода. Через входной порт агент получает сообщения и запросы от других агентов, а также информацию из других источников, не являющихся агентами в указанном смысле. Через выходной порт агент выдает сообщения и запросы другим агентам, а также информацию получателям, не являющимися агентами.

В основе, рассматриваемого в данной статье подхода лежит идея о разграничении синтаксического и семантического описания свойств образов на языке ситуационного модального нечеткого исчисления [5]. Синтаксическое описание служит средством обмена между агентами и представляет онтологию на формальном языке ситуационного модального нечеткого исчисления. Семантическое описание используется агентами для вычислений и представляет онтологию на языке нечетких отношений. Его можно рассматривать как своеобразную формальную систему со своим языком, правилами вывода и интерпретацией. Вывод заключений делается на семантическом уровне. Связь между синтаксическим и семантическим уровнями осуществляется посредством двух процедур: процедуры семантической трансляции, которая выполняет перевод текста, представленного на синтаксическом уровне на язык нечетких отношений, то есть на семантический уровень; процедуры лингвистической аппроксимации, которая выполняет перевод заключений, полученных на семантическом уровне, на язык синтаксического уровня. Агенты обмениваются информацией на синтаксическом уровне.

Запрос к БДУВ является теоремой о наличии у образа некоторых свойств. Свойства образа представляются правилами нечеткого исчисления $p(G_F)$, $\varphi(B(N_j))$, заданными на универсумах U . Эти свойства образуют теорию для осуществления вывода (ответа на запрос). Установление наличия этих свойств осуществляется с помощью процедуры вывода (доказательства) в нечетком исчислении. Подробно исчисление и процедура вывода описаны в [5]. В процессе вывода агенты, отвечающие за опрос датчиков, помещают в БДУВ по мере необходимости значения параметров образов, необходимых для вывода. В процессе вывода могут участвовать несколько агентов, обменивающихся друг с другом сообщениями.

3 Пример

Рассмотрим простой иллюстративный пример. Этот же пример, но несколько в другой интерпретации рассматривается в [5]. Точечный объект находится в некотором месте N_i двумерного пространства. Координаты объекта изменяются в определенных пределах по осям абсцисс и ординат. В соответствии с предыдущими разделами, это можно интерпретировать как наличие единственного множества $B(N_i)$, $i = 1$, определяющего местоположение точечного объекта и отсутствие правил $p(GF)$. Будем характеризовать положение объекта по осям соответственно лингвистическими переменными: $\langle X, \Theta_B(X), U, Z, M \rangle$ и $\langle Y, \Theta_B(Y), V, Z, M \rangle$, где: X и Y – лингвистические переменные; характеризующие положение объекта соответственно по осям абсцисс и ординат; $\Theta_B(X) = \Theta_B(Y) = \{\text{Слева}, \text{В_центре}, \text{Справа}\}$ – базовые множества лингвистических значений; $U = V = [-1, 1]$ – универсумы значений координат по осям X и Y ; Z – синтаксическое правило, M семантическое правило. Базовые множества лингвистических значений обоих лингвистических переменных выбраны для удобства изложения совпадающими. Одинаковые размерности универсумов значений координат от -1 до 1 по обеим осям также для удобства изложения выбраны совпадающими. Будем полагать, что $U=V=\{-1, -0.8, -0.6, -0.4, -0.2, 0, 0.2, 0.4, 0.6, 0.8, 1\}$.

Смысловые содержания лингвистических значений *Слева*, *В_центре*, *Справа* определяют семантическими правилами: $M(\text{Слева}) = \mu_L$, $M(\text{В_центре}) = \mu_C$, $M(\text{Справа}) = \mu_R$. Графики функций принадлежности μ_L , μ_C , μ_R представлены на рисунке 5.

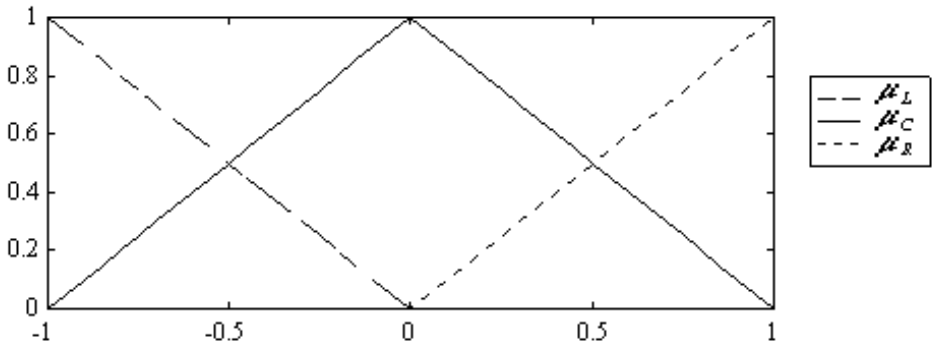


Рисунок 5 – Графики функций принадлежности.

Онтология $\varphi(B(N_i))$, $i = 1$)содержит следующие аксиомы, выражающие свойства расположения вершин (координат) $B(N_i)$:

- (1) (X есть Слева \supset Y есть Справа)
- (2) (X есть В_центре \supset Y есть В_центре)
- (3) (X есть Справа \supset Y есть Слева)
- (4) (Возможно_50 X есть В_центре & Возможно_50 X есть Справа)

Требуется найти ответ (лингвистическую аппроксимацию) на вопрос, содержательная суть которого следующая: «Если известна онтология $\varphi(B(N_i))$, подставленная аксиомами (1)-(4), то какова неизвестная координата объекта?». Аксиома 4 определяет координату объекта по оси X. Поэтому искать следует только координату объекта по оси Y . В исчислении вопрос задается следующим выражением

- (5) (m_1 Y есть Слева & (m_2 Y есть В_центре & m_3 Y есть Справа))?

В этом выражении надо определить значения модальностей m_1, m_2, m_3 , характеризующие расположение объекта по оси Y.

Проведем синтаксическую трансляцию онтологии (1)-(4), а затем лингвистическую аппроксимацию выражения (5).

Имеем две лингвистические переменные X, Y. Для этого набора лингвистических переменных возможны три основы схем отношений $\{X\}$, $\{Y\}$, $\{X, Y\}$. Подробно проиллюстрируем семантическую трансляцию первой аксиомы онтологии: (X есть Слева \supset Y есть Справа). В данную формулу входят атомы X есть Слева и Y есть Справа. Результатом семантической трансляции этих атомов являются нечеткие отношения $R_L[X] = \langle \{X\}, M(\text{Слева}) \rangle$ и $R_R[Y] = \langle \{Y\}, M(\text{Справа}) \rangle$, где $R_L = M(\text{Слева})$ – нечеткое множество в универсуме $U = [-1, 1]$, $R_R = M(\text{Справа})$ – нечеткое множество в универсуме $V = [-1, 1]$. Согласно рисунку 5 нечеткие множества R_L и R_R имеют функции принадлежности, представленные соответственно в таблицах 1 и 2.

Таблица 1 – Функция принадлежности множества R_L .

x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
$\mu_{R_L}(x)$	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0

Таблица 2 – Функция принадлежности множества R_R .

y	$\mu_{R_R}(y)$
-1	0
-0.8	0
-0.6	0
-0.4	0
-0.2	0
0	0
0.2	0.2
0.4	0.4
0.6	0.6
0.8	0.8
1	1

Результатом семантической трансляции формулы (X есть Слева $\supset Y$ есть Справа) является нечеткое отношение $R_I[X, Y] = C_{\{Y\}} R_L[X] \supset C_{\{X\}} R_R[Y]$, где R_I - нечеткое множество в g

$V = [-1, 1] \times [-1, 1]$. Для вычисления его функции принадлежности необходимо получить цилиндрическое продолжение нечеткого отношения $R_L[X]$ на основу схемы $\{Y\}: C_{\{Y\}} R_L[X]$ и цилиндрическое продолжение нечеткого отношения $R_R[Y]$ на основу схемы $\{X\}: C_{\{X\}} R_R[Y]$. Функции принадлежности этих цилиндрических продолжений приведены в таблицах 3 и 4 соответственно. Они представлены в виде матриц, в которых строки помечены элементами универсума V , а столбцы – элементами универсума U . На пересечении столбца x и строки y находятся значения функции принадлежности $\mu(x, y)$.

Следующий шаг вычисления функции принадлежности нечеткого отношения $R_I[X, Y]$ – применение операции следования к полученным цилиндрическим продолжениям. Получившаяся в итоге функция принадлежности $\mu_{R_I}(x, y)$ приведена в матричной форме в таблице 5.

Семантическую трансляцию остальных аксиом онтологии так подробно проводить не будем, а приведем лишь обозначения семантической трансляции атомов (таблица 6) и выражения, по которым вычисляются результирующие нечеткие отношения (таблица 8). В таблицах 7 и 8 в левом столбце указана формула, а в правом – результат ее семантической трансляции.

Таблица 3 – Функция принадлежности цилиндрического продолжения $C_{\{Y\}} R_L[X]$.

$y \setminus x$	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
-1	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
-0.8	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
-0.6	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
-0.4	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
-0.2	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
0	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
0.2	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
0.4	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
0.6	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
0.8	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0
1	1	0.8	0.6	0.4	0.2	0	0	0	0	0	0

Таблица 4 - Функция принадлежности цилиндрического продолжения $C_{\{X\}}R_R[Y]$.

y \ x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
-1	0	0	0	0	0	0	0	0	0	0	0
-0.8	0	0	0	0	0	0	0	0	0	0	0
-0.6	0	0	0	0	0	0	0	0	0	0	0
-0.4	0	0	0	0	0	0	0	0	0	0	0
-0.2	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
1	1	1	1	1	1	1	1	1	1	1	1

Таблица 5 – Функция принадлежности отношения $R_I[X, Y]$.

y \ x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
-1	0	0	0	0	0	1	1	1	1	1	1
-0.8	0	0	0	0	0	1	1	1	1	1	1
-0.6	0	0	0	0	0	1	1	1	1	1	1
-0.4	0	0	0	0	0	1	1	1	1	1	1
-0.2	0	0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	0	1	1	1	1	1	1
0.2	0.2	0.2	0.2	0.2	1	1	1	1	1	1	1
0.4	0.4	0.4	0.4	1	1	1	1	1	1	1	1
0.6	0.6	0.6	1	1	1	1	1	1	1	1	1
0.8	0.8	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1

Семантическую трансляцию остальных аксиом онтологии так подробно проводить не будем, а приведем лишь обозначения семантической трансляции атомов (таблица 6) и выражения, по которым вычисляются результирующие нечеткие отношения (таблица 7). В таблицах 6 и 7 в левом столбце указана формула, а в правом – результат ее семантической трансляции.

Таблица 6 – Обозначения семантической трансляции атомов.

Формула языка	Нечеткое отношение
<i>X есть Слева</i>	$R_L[X] = \langle \{X\}, M(\text{Слева}) \rangle$
<i>Y есть Справа</i>	$R_R[Y] = \langle \{Y\}, M(\text{Справа}) \rangle$
<i>X есть В_центре</i>	$R_C[X] = \langle \{X\}, M(\text{Слева}) \rangle$
<i>Y есть В_центре</i>	$R_C[Y] = \langle \{Y\}, M(\text{В_центре}) \rangle$
<i>X есть Справа</i>	$R_R[X] = \langle \{X\}, M(\text{Справа}) \rangle$
<i>Y есть Слева</i>	$R_L[Y] = \langle \{Y\}, M(\text{Слева}) \rangle$

Таблица 7 – Выражения для вычисления нечетких отношений.

Формула языка	Нечеткое отношение
$(X \text{ есть Слева} \supset Y \text{ есть Справа})$	$R_1[X, Y] = C_{\{Y\}} R_L[X] \supset C_{\{X\}} R_R[Y]$ g
$(X \text{ есть В_центре} \supset Y \text{ есть В_центре})$	$R_2[X, Y] = C_{\{Y\}} R_C[X] \supset C_{\{X\}} R_C[Y]$ g
$(X \text{ есть Справа} \supset Y \text{ есть Слева})$	$R_3[X, Y] = C_{\{Y\}} R_R[X] \supset C_{\{X\}} R_L[Y]$ g
$(\text{Возможно}_{50} X \text{ есть В_центре} \& \text{Возможно}_{50} X \text{ есть Справа})$	$R_4[X] = \lambda_{(2,0)} R_C[X] \cap \lambda_{(2,0)} R_R[X]$

В табл. 8 и 9 приведены окончательные результаты семантической трансляции второй и третьей аксиомы: функции принадлежности нечетких отношений $R_2[X, Y]$ и $R_3[X, Y]$, представленные в матричной форме.

Таблица 8 -Функция принадлежности отношения $R_2[X, Y]$.

y \ x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
-1	1	0	0	0	0	0	0	0	0	0	1
-0.8	1	1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	1	1
-0.6	1	1	1	0.4	0.4	0.4	0.4	0.4	1	1	1
-0.4	1	1	1	1	0.6	0.6	0.6	1	1	1	1
-0.2	1	1	1	1	1	0.8	1	1	1	1	1
0	1	1	1	1	1	1	1	1	1	1	1
0.2	1	1	1	1	1	0.8	1	1	1	1	1
0.4	1	1	1	1	0.6	0.6	0.6	1	1	1	1
0.6	1	1	1	0.4	0.4	0.4	0.4	0.4	1	1	1
0.8	1	1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	1	1
1	1	0	0	0	0	0	0	0	0	0	1

Таблица 9 – Функция принадлежности отношения $R_3[X, Y]$.

y \ x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
-1	1	1	1	1	1	1	1	1	1	1	1
-0.8	1	1	1	1	1	1	1	1	1	1	0.8
-0.6	1	1	1	1	1	1	1	1	1	0.6	0.6
-0.4	1	1	1	1	1	1	1	1	0.4	0.4	0.4
-0.2	1	1	1	1	1	1	1	0.2	0.2	0.2	0.2
0	1	1	1	1	1	1	0	0	0	0	0
0.2	1	1	1	1	1	1	0	0	0	0	0
0.4	1	1	1	1	1	1	0	0	0	0	0
0.6	1	1	1	1	1	1	0	0	0	0	0
0.8	1	1	1	1	1	1	0	0	0	0	0
1	1	1	1	1	1	1	0	0	0	0	0

Особо следует выделить семантическую трансляцию последней аксиомы онтологии, так как она иллюстрирует применение операции линейного преобразования. В таблице 10 приве-

дены результаты семантической трансляции аксиомы *Возможно_50 X есть В_центре & Возможно_50 X есть Справа*.

Таблица 10 – Результаты семантической трансляции аксиомы (4).

x	-1	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1
$\mu_{R_C}(x)$	0	0.2	0.4	0.6	0.8	1	0.8	0.6	0.4	0.2	0
$\lambda_{(2,0)} \mu_{R_C}(x)$	0	0.4	0.8	1	1	1	1	1	0.8	0.4	0
$\mu_{R_R}(x)$	0	0	0	0	0	0	0.2	0.4	0.6	0.8	1
$\lambda_{(2,0)} \mu_{R_R}(x)$	0	0	0	0	0	0	0.4	0.8	1	1	1
$\mu_{R_4}(x)$	0	0	0	0	0	0	0.4	0.8	0.8	0.4	0

Таким образом, синтаксическая трансляция аксиом завершена и можно переходить к выводу заключения.

- | | |
|---|------------------------------------|
| 1) $R_1[X, Y]$, | аксиома, |
| 2) $R_2[X, Y]$, | аксиома, |
| 3) $R_3[X, Y]$, | аксиома, |
| 4) $R_4[X]$, | аксиома, |
| 5) $R_5[X, Y] = C_{\{Y\}} R_4[X]$, | из 4) по правилу обобщения, |
| 6) $R_6[X, Y] = R_1[X, Y] \cap R_2[X, Y]$, | из 1) и 2) по правилу пересечения, |
| 7) $R_7[X, Y] = R_6[X, Y] \cap R_3[X, Y]$, | из 6) и 3) по правилу пересечения, |
| 8) $R_8[X, Y] = R_7[X, Y] \cap R_5[X, Y]$, | из 7) и 5) по правилу пересечения, |
| 9) $R_9[Y] = Proj_{\{Y\}} R_8[X, Y]$, | из 8) по правилу проекции. |

Опуская все промежуточные вычисления (их можно посмотреть в [5]) перейдем к главной цели – выводу заключения (ответу на вопрос) как лингвистической аппроксимации нечеткого отношения $R_9[Y]$ в виде формулы ($m_1 Y$ есть Слева & ($m_2 Y$ есть В_центре & $m_3 Y$ есть Справа)), в которой нам предстоит определить символы модальности m_1 , m_2 и m_3 . Для этого необходимо вычислить степени совместимости $\tau_1 = CP(R_L[Y], R_9[Y])$, $\tau_2 = CP(R_C[Y], R_9[Y])$ и $\tau_3 = CP(R_R[Y], R_9[Y])$.

Для определения значений функции принадлежности степени совместимости можно воспользоваться следующей методикой. Например, для того, чтобы определить значение функции принадлежности степени совместимости $\tau_2 = CP(R_C[Y], R_9[Y])$ при $x=0.4$, необходимо вычислить, при каких значениях аргумента $\mu_C(y)=0.4$. В данном случае имеем $y \in \{-0.6, 0.6\}$. Далее, из значений $\mu_{R_9}(-0.6)=0.8$ и $\mu_{R_9}(0.6)=0$ следует выбрать наибольшее, в данном случае 0.8 . Это число и будет значением $\mu_{\tau_2}(0.4)$.

Результатом лингвистической аппроксимации нечеткого отношения $R_9[Y]$ является формула (*Возможно_50 Y есть Слева & (Возможно_50 Y есть В_центре & Невозможно_95 Y есть Справа)*), которую можно интерпретировать как "Y - левее центра".

Заключение

Распределенные беспроводные и комбинированные сети интеллектуальных датчиков и связанные с ними системы сбора, обработки и передачи информации становятся новой областью исследований. В предлагаемой статье были рассмотрены принципы представления и использования информации собираемой с датчиков для анализа свойств образов (ответа на запросы о свойствах). Свойства образов, описываемые в этом исчислении (онтология) состоят из последовательностных (т.е. наблюдаемых во времени и/или пространстве) свойств образов,

мгновенных свойств, задаваемых в определенных точках времени или пространства и фактических значений измеряемых параметров (универсумов значений). Онтология хранится не обязательно централизованно. Она может быть распределена по агентам, каждый из которых отвечает за свою часть анализа свойств образов. Для формирования последовательных свойств предлагается подход, основанный на синтаксическом анализе структуры образов в n-мерном пространстве. В качестве аппарата описания всех совокупных свойств образов используется нечеткое ситуационное исчисление. Все свойства представляются на языке нечеткого ситуационного исчисления как теория, а анализ образов является выводом в этой теории.

Настоящая работа частично выполняется по теме «Интегрированный телекоммуникационный узел сбора, обработки и анализа телеметрической информации для удаленного представления результатов анализа службам городского хозяйства» в рамках государственного контракта на выполнение работ по федеральной целевой программе «Интеграция науки и высшего образования России на 2002-2006 годы». Экспериментальной базой являются сети сбора информации с распределенных датчиков, передачи снимаемых с них данных по мобильной или проводной связи в узел беспроводной связи, последующая передача, используя Интернет, данных для анализа свойств образов на сервер анализа запросов. Для экспериментального апробирования используются несколько инструментальных пакетов: Visual Prolog, AgentBuilder и ряд других.

Список литературы

- [1] Estrin D., Govindan R., Heidemann J., Kumar S. Next century challenges: Scalable coordination in sensor networks // ASM MOBICOM. 1999.
- [2] Bonnet P., Gehrke J.E., Seshadri P. Towards sensor database systems // ASM Mobile Data Management. 2001.
- [3] Takayama S., Kariya K. Autonomous measuring by sensing node in telemetry system // Measurement science review. Vol. 3, section 3. 2003. Pp. 29-32.
- [4] Девятков В.В. Онтологии и их применение // Программные продукты и системы. 2000. 3.
- [5] Девятков В.В. Мультиагентное иерархическое распознавание на основе нечеткого ситуационного исчисления. Труды ИГУ РАН, том XX, 2003. С. 78 -99.
- [6] Russel S.J., Norvig P. Artificial Intelligence. A modern Approach. - NJ, Prentice-Hall, 1998.

МОЖЕТ ЛИ КОМПЬЮТЕР БЫТЬ СУБЪЕКТОМ ПРАВА?

С.В. Тиминский

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
timinsky_s@mail.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: правоспособность, дееспособность, субъект права, юридическое лицо, интеллектуальные агенты

Abstract

This paper concerns the possibility of acknowledging the legal personality by the software intelligent agents. As mediators and as initiators of electronic transactions, intelligent agents may truly facilitate electronic commerce by fostering a global and digital economy. As mentioned below, there are a number of ways in which a confidence boost may be provided such as by developing a certification system for agents, or by establishing a labeling system for them. This is likely to be the main issue on which the extent of use of intelligent agent systems in the future will depend.

Введение

Наверное, очень трудно даже представить в своем воображении судебный процесс, истцом или ответчиком по которому будет проходить компьютерная программа. На взгляд непосвященного, такое выглядит невозможным, также как невозможно, к примеру, судиться с миксером, кофемолкой или вообще с любой другой вещью. У вещи не может быть прав, обязанностей, поэтому у ней не может быть имущества на праве собственности, на которое может быть обращено взыскание; наконец, это просто кажется лишенным всякого смысла, признать вещь субъектом права, поскольку это равносильно утверждению за вещь способности адекватно воспринимать реальность, разумно мыслить и действовать – что воспринимается как не совсем удачная шутка.

Тем не менее, вопрос правоспособности компьютера не обходится стороной в зарубежных исследованиях по юриспруденции [1-5], чего нельзя сказать о правовой науке нашего отечества. Современный же уровень и скорость развития информационных технологий делает эту проблему все более актуальной, а экскурс в область истории и теории государства и права высвечивает довольно интересные ее особенности.

1 Интеллектуальные агенты и заключение договора

Как пишут Т. Аллен и Р. Уиддисон, “появляющиеся в настоящее время компьютерные системы могут действовать не просто автоматически, но автономно. Автономные машины способны обучаться на собственном опыте, изменяя инструкции в их собственных программах и даже создавая для себя новые инструкции. Они могут принимать решения, основанные на этих ими же измененных или ими же созданных инструкциях... Эти процессы включают осуществление выбора, формирование намерений, принятие решений и предоставление либо непредоставление согласия” [1].

Такого рода компьютерные программы получили в технической литературе название *интеллектуальных агентов*. В широком смысле понятие интеллектуальных агентов включает в себя следующие свойства:

- “автономность – агенты действуют без прямого вмешательства людей или других агентов и способны в определенной степени контролировать свои действия и свое внутреннее состояние;
- коммуникативность – агенты взаимодействуют с другими агентами (и, возможно, также с людьми) посредством некоторого рода *языка для коммуникации агентов*;
- реагирование – агенты изучают окружающую их среду (это может быть физический мир, пользователь – через графический пользовательский интерфейс, другие агенты, интернет, или, возможно, все сочетания перечисленного) и своевременно отвечают на изменения, которые в ней происходят;
- профессиональная активность – агенты не просто действуют при изменении их окружающей среды, они способны демонстрировать целенаправленное поведение посредством *взятия инициативы*” [2].

В узком смысле помимо перечисленных характеристик понятие ‘интеллектуальный агент’ включает в себя одну или несколько из следующих особенностей:

- “мобильность – способность агента перемещаться в электронном сетевом окружении;
- достоверность и точность – предположение, что агент не будет использовать информацию, если будет знать, что она является ложной;
- благожелательность – предположение, что агенты не будут иметь противоречивых целей;
- рациональность – предположение, что агент будет действовать так, чтобы достичь своих целей и не будет действовать так, чтобы предотвратить их достижение, хотя бы в той степени, в которой позволяют его представления об этом” [2].

Появление компьютерных программ с подобными свойствами вызовет определенные проблемы в области права – в частности, в сфере возмещения ущерба и в договорном праве. Как отмечает К. Карноу “результатом некоторых решений [интеллектуальных агентов] будет “патология”; то есть решение будет не только непредсказуемым или удивительным, что может быть вполне желательным, но также может иметь непредусмотренные и разрушительные последствия как для сетевого пространства, так и для значительной части всей нашей инфраструктуры, подсоединенной к такому сетевому пространству”[3]. Должен ли человек нести ответственность за действия, автономного компьютера, который способен на принятие самостоятельных решений? Ряд исследователей полагает, что нет. Л. Уэйн, например, пишет: “оставленные без контроля интеллектуальные артефакты должны нести ответственность независимо от людей ими владеющих, с позиции универсальной концепции ответственности основных правовых систем в течение истории”[4].

Кроме того, внедрение интеллектуальных компьютерных агентов в электронную коммерцию приводит к появлению такого типа договоров, заключение которых с одной или с обеих сторон происходит без непосредственного участия человека, и даже более того, человек, в интересах которого компьютерная программа заключает договор, не знает о его существовании вплоть до стадии его исполнения [1, 5]. С позиции традиционной договорной доктрины – как в системе общего праве, так и континентального – в действительности таких договоров возникают очень большие сомнения ввиду отсутствия критерия “совпадения воли”, необходимого для возникновения любого соглашения.

2 Восток и Запад

В некоторой степени может быть справедливо утверждение, что аналог Интернета – «локальные, территориальные и глобальные сети» появился в СССР раньше, чем в США [10]. Тем не менее, существование в СССР плановой экономики не лучшим образом повлияло на потребительский рынок – нормой был дефицит товаров и большие очереди.

В США же во время холодной войны процветала рыночная экономика, большое развитие получили так называемые удаленные или дистанционные контракты (*distant contracts*) – теле-

магазин или система рассылок товарных каталогов (советским аналогом была система посыл-торга) и контракты, заключаемые путем покупки товара у автомата. В СССР автоматические покупки не были сильно развиты – исключение составляли лишь автоматы с газированной водой. Творчество юристов лишь отражает состояние экономики того или иного государства, поэтому западная юриспруденция знает немало работ, посвященных заключению дистанционных или автоматических контрактов, а советская и российская правовая наука этих вопросов едва касалась.

Переход РФ на рыночную экономику в 90-х годах привел к возникновению такой ситуации, что купля-продажа путем дистанционных (магазин на диване, интернет-магазин) контрактов стала развиваться. Все большее число бизнес-процессов, хотя и медленно (предпочтение пока что отдается факсам, а не электронной почте) автоматизируются, человек делегирует свой труд компьютерному роботу. В том числе фактически компьютерной программе передаются полномочия по заключению, изменению и расторжению договора с потребителем или коммерсантом. Поэтому приобретает особую актуальность вопрос о том, не лучше ли в интересах удобства придать такой программе-роботу статус субъекта права, чтобы освободить человека от риска нести ответственность за ошибки или неточности, которые допускает робот при работе с контрактами.

3 Правоспособность и дееспособность. Исторический экскурс

Правоспособность есть способность субъекта иметь права и обязанности. По современному правосознанию каждый человек правоспособен прежде всего потому, что он человек, не составляют исключение даже малолетние дети и умалишенные лица. Хотя эти лица фактически лишены возможности проявлять юридически признаваемую волю, тем не менее они считаются способными обладать юридически обеспеченной сферой власти. Право признат за ними способность вырабатывать волю; оно считает их только временно лишенными этой способности, предполагая, что даже неизлечимо больные могут выздороветь [6]. Пока эти лица не имеют разумной воли, право восполняет этот недостаток при помощи института законных представителей.

Юридические отношения возникают главным образом путем действий лиц, участвующих в гражданском обороте. Действие в юридическом смысле есть внешнее проявление акта воли. Но с юридической точки зрения, лишь тот, кто может принимать решения, взвешивая мотивы и принимая во внимание внешние последствия действий. Поэтому дети и душевнобольные люди не могут *действовать* в юридическом смысле слова. За лицами, которые *могут* действовать, право признает способность вызывать путем своих действий известные юридические последствия. Эта способность называется дееспособностью. В работе [6] дееспособность подразделяется на способность к совершению деликтов (правонарушений) – вменяемость или деликтоспособность – и способность к совершению юридических сделок.

История права знает немало примеров, когда правосубъектностью наделялись даже неодушевленные предметы и, наоборот, когда люди не рассматривались в качестве субъектов права, а отсутствие у них *persona*, необходимого для участия в гражданском обороте, приравнивало их к неодушевленным вещам или животным.

Так, Дж. Ч. Грэй пишет, что храмы в Древнем Риме и здания церкви в Средние века пользовались отдельным корпоративным бытием в качестве субъектов права. Грэй также замечает: “Неодушевленные вещи рассматривались как носители законных обязанностей, - я собирался добавить, в первобытные времена, но, как мы увидим, понятие это упорно сохранялось даже в наши дни. Если здесь имела фикция, она не заключалась в том, чтобы приписать вещи истинную волю человека, но в предположении, что вещь имеет свой собственный ум. Мне кажется, однако, что часто не было никакой сознательной фикции, но была некая определенно не осознанная вера в то, что вещь действительно обладает умом и волей” [12].

Тем не менее, как отмечает Л. Уэйн, с исторической точки зрения, “изошренная конструкция вещи как полномочного представителя человека уходит своими корнями не в отношения между людьми и неодушевленными объектами, но в отношения между одними людьми и другими” [4] – а именно, как это ни парадоксально, в институт рабства¹.

Раб в древнем Риме находился вне политического общества и не являлся субъектом ни *ius civile*, ни *ius gentium*; по римскому праву, как и в других античных правовых системах, он считался вещью (*res*) [7], говорящим орудием (*instrumentum vocatio*) [8]. В контексте нашей тематики сущность института рабства заключалась не только в том, что право полагало человека вещью, но еще и в том, что эта вещь постепенно стала допускаться к участию в гражданском обороте, оставаясь при этом “в глазах” права вещью на протяжении очень долгового времени². Правовой инструментарий, применяемый для этого древнеримскими юристами, был весьма оригинален:

Первоначально участие раба в коммерческом обороте подчинялось принципу, что заключенные им сделки могут только улучшать положение его господина, но не ухудшать. И.Б. Новицкий комментирует это так: “Разумеется, такое примитивное построение - по сделкам раба права приобретаются господином, а обязанности ложатся на раба (с которого ввиду его неспособности получить было нельзя) – не могло сохраниться с развитием торговли и с усложнением хозяйственной жизни. Желających вступать в сделки с рабами при полной безответственности по этим сделкам самого рабовладельца нашлось бы немного. Правильно понявший интерес рабовладельца требовал, чтобы третьи лица, с которыми вступал в деловые отношения раб, могли рассчитывать на возможность осуществления своих прав по сделкам с рабами”.

В интересах рабовладельца происходило постепенное признание личности раба [8]. На этой почве сложился институт *пекулия*. Термином «пекулий» называлось имущество, выделяемое из общего имущества рабовладельца в управление раба. Поскольку управлять имуществом невозможно без совершения различных сделок, то не признавая раба правоспособным лицом, древние римляне признали, однако, юридическую силу за совершаемыми им сделками.

Далее, современное английское право в определенной степени распространяет правосубъектность на корабли с целью обеспечения интересов кредиторов. Это позволяет заинтересованным лицам выдвигать иски непосредственно против судна, ... указав его в качестве ответчика. Как отмечают Т. Аллен и Р. Уиддисон, в данном случае надделение кораблей некой разновидностью правосубъектности представляет собой достижение определенно значимой для права цели удобным и относительно недорогим способом [1].

4 Лицо и персональность. Юридическое лицо

Современная гражданско-правовая наука воспринимает любого носителя прав и обязанностей со значительной долей абстракции, обращая внимание лишь на внешних социально-значимых его особенностях. Идея абстрактного субъекта правоотношения получила соответствующее воплощение в понятии *лица*: латинский термин *persona* (лицо) первоначально означал ни что иное как *театральную маску*. С позиции права все *лица* - это всего лишь *маски* с некоторым набором прав и обязанностей в зависимости от степени участия в гражданском обороте.

Негативным эффектом такого положения дел является существование теневой экономики с так называемыми фирмами-однодневками, существующими только номинально, с единственной целью ухода от налогообложения стоящих за ними *реальных* субъектов.

¹ Еще в 1950 году Н. Уинер писал: “автоматические механизмы, что бы мы ни думали по поводу тех чувств, которые они испытывают или не испытывают, являются точным экономическим эквивалентом рабского труда”- N. Wiener, *The Human Use Of Human Beings*, p. 189 (1950), цит. по [4].

² Рабы стали допускаться до участия в гражданском процессе лишь после соответствующего рескрипта императора Юстиниана (VI в. н.э.) – см. [7]

Позитивным же следствием персональной абстракции является то, что «под понятие субъектов права ... подходят, кроме отдельных людей, еще те *социальные организмы* которые также преследуют *человеческие интересы* и имеют особую организацию, позволяющую им выработать самостоятельную волю, отличную от воли отдельных сочленов, входящих в состав этих социальных организмов, и стоящую над волей этих последних» [6]. Эти коллективные субъекты получили в современной теории права название *юридических лиц*.

Сторонники признания правоспособности компьютера в качестве исторической аналогии приводят в пример именно конструкцию юридического лица – которое долгое время считалось (а в англо-американском праве считается до сих пор – см. [9]) искусственной конструкцией (теории фикции). Родоначальником этого подхода принято считать римского папу Иннокентия IV. В 1245 г. на вопрос о возможности отлучения корпорации от церкви он заявил, что корпорация не имеет души, а существует лишь в воображении людей, будучи *persona ficta* или *corpus mysticum*, т.е. фиктивным, не существующим в реальности лицом [13].

Таким образом, при решении вопроса о правосубъектности компьютера можно не принимать во внимания такие факторы, как наличие или отсутствие у компьютера внутренней воли, самосознания и других подобных качеств, хотя некоторые исследователи эту проблему затрагивают, выдвигая гипотезу о существовании некоего морального права для признания любого субъекта права со стороны правовой системы [15].

Теория социальной реальности утверждает, что коллективные группы могут развиваться в общественном организме с различной социальной волей. Правовая система просто распознает социальный факт их независимого существования, наделяя их правосубъектностью. Как заметил Тебнер: «... как правило, стало прововой политикой наделять правосубъектностью социальные системы, которые уже обладали социальной способностью для коллективного действия, [...] правовая система подверглась огромному воздействию, чтобы завершить социосубъектность правосубъектностью» [10].

Юридические лица являются социальными системами, в то время как компьютеры – это информационные системы. Но есть одно ключевое сходство: в случае с организациями, проблема заключается в отграничении сущности и действий юридического лица от сущности и действий его учредителей. Другими словами, способность к коллективным действиям в обществе составляет ее сущность. В случае с компьютерами, проблема также заключается в отграничении сущности и действий компьютера от людей, осуществляющих над ним контроль [1]. Поэтому можно трансформировать утверждение Тебнера следующим образом: «становится смыслом правовой политики предоставление правосубъектности *информационным* системам, которые уже имеют способность к *автономным* действиям».

По аналогии с юридическими лицами предлагается решить и проблему идентификации. Одно решение, предложенное скандинавскими правоведами [11] – это внедрение различных классификаций по надежности (*security classification*) и сертификация интеллектуальных агентов со ссылкой на соответствующий класс стандартов надежности. При этом предъявляются различные требования в отношении уровней надежности, которым интеллектуальные агенты должны соответствовать в том случае, если они допускаются к определенной деятельности. Такая система может потребовать контроля, определяющего соответствие агентом специфическому уровню надежности. Это может привести, в свою очередь, к развитию системы независимых отметок о проверке (*verification marks*) для особенностей надежности агента.

Сходную систему маркировки (*labelling*) агентов предлагает использовать Ж.-Ф. Леруж [5], отличием выступает ее добровольный, а не принудительный характер, что обеспечивало бы большую гибкость.

К. Карноу предложил систему сертификации электронных агентов, благодаря которой агенты смогут использоваться только после их сертификации. Сертификационный орган (Регистр Тьюринга) оценит риски, связанные с использованием агента, и соответственно этой

оценке он выдаст сертификат для агента. Такой агент будет способен функционировать только с системами, которые также имеют сертификат, выпущенный Реестром [3].

Заключение

Как посредники и как инициаторы сделок, заключаемых и/или исполняемых в информационно-коммуникационном пространстве интеллектуальные агенты могут действительно упростить электронную коммерцию, принимая участие в создании глобальной и цифровой экономики. Однако, для того, чтобы это произошло в действительности, как предприниматели, так и потребители должны сначала почувствовать удобство феномена интеллектуальных систем. Очень часто пользователь не только не видит, что на самом деле происходит в информационной системе с его правами и обязанностями, но и не знает этого. Подобные ситуации не приводят к чувствам комфорта и безопасности, необходимых для активного использования таких систем в различных общественных сферах.

Как упоминалось выше, существует ряд способов, когда рост доверия может быть вызван развитием особой сертификационной системы для интеллектуальных агентов или введением системы для их маркировки. Это, по всей видимости, будет основным вопросом, от которого будет зависеть степень использования систем интеллектуальных агентов.

Список литературы

- [1] Tom Allen & Robin Widdison, *Can computers make contracts?*, 9 Harvard Journal of Law and Technology, 25 (1996).
- [2] M. Wooldridge & N. Jennings, *Intelligent Agents: Theory and Practice*, Knowledge Engineering Review Vol. 10 No. 2, June 1995, Cambridge University Press, 1995.
- [3] С.Е.А. Karnow *Future Codes: Essays in Advanced Computer Technology and the Law*, Artech House, 1997.
- [4] L.E. Wein, *The responsibility of intelligent artifacts: Toward an automation jurisprudence*, 6 Harvard Journal of Law & Technology, 109-110 (1992).
- [5] J.F. Lerouge, "The use of electronic agents questioned under contractual law: Suggested solutions on a European and American Level", *The John Marshall Journal of Computer & Information Law*, Vol. XVIII No. 2, Winter 2000, pp. 403 et seq.
- [6] Хвостов В.М. *Система римского права*, М.: Спарк, 1996.
- [7] Дождев Д.В. *Римское частное право*, М., 1996.
- [8] Новицкий И.Б. *Римское право*, М., 1994.
- [9] Батлер У.Э., Гаши-Батлер М.Е. *Корпорации и ценные бумаги по праву России и США*, М., 1997, с.21
- [10] *Biological Metaphors in Legal Thought, Rottleuthner in Autopoietic Law: A New Approach to Law and Society*, Teubner (ed.) (1988, Walter de Gruyter, Berlin/New York)
- [11] Юзвизин И.И. *Основы информатиологии*. Учебник. – М., 2000, с.442 и далее.
- [12] J.C. Gray, *The Nature And Sources Of The Law*, (2d ed. 1921)
- [13] Венедиктов А.В. *Государственная социалистическая собственность*, Л., 1948, с. 675
- [14] Stuurman, C. & Wijnands, H.S.A. (2001). *Intelligent agents: a curse or a blessing: a survey of the legal aspects of the application of intelligent software systems*. The Computer Law and Security Report, 17(2), 92-99.
- [15] L. Solum, *Legal Personhood for Artificial Intelligences* (1992) 70 North Carolina Law Review p 1231.

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ УПРАВЛЕНИЯ РЕГИОНОМ И РЕШЕНИЕ ЗАДАЧ ФЦП «ЭЛЕКТРОННАЯ РОССИЯ 2002-2010 ГГ.» С ПОМОЩЬЮ МУЛЬТИАГЕНТНЫХ ТЕХНОЛОГИЙ

В.А. Виттих¹, Г.Д. Светкина², П.О. Скобелев³, Е.А. Гриценко², О.Л. Сурнин⁴, Д.В. Волхонцев⁴

¹Институт проблем управления сложными системами РАН
443020, Россия, Самара, ул. Садовая, 61
vittikh@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²Министерство труда и социального развития Самарской области
Россия, Самара, ул. Революционная, 44
тел: +7 (8462) 70-66-84, факс: +7 (8462) 70-66-85

³ООО «НПК «Маджента Девелопмент»
443010, Россия, Самара, ул. Осипенко, 1
skobelev@magenta-technology.ru

тел: +7 (8462) 70-66-84, факс: +7 (8462) 70-66-85

⁴ООО «НПК «Генезис знаний»

443010, Россия, Самара, ул. Ново-Садовая, 14а
surnin@kg.ru

Ключевые слова: эффективное управление регионом, внедрение мультиагентных разработок, мультиагентные технологии, взаимодействие власти и населения, информационные технологии в управлении

Abstract

This article describes the problems of citizen to government interaction when solving social problems. The article sets goals and tasks for creation of the system for region management with the use of multi-agent technologies in the Samara Region. The authors give comparative analysis of the goals of the Federal program “Electronic Russia” and goals of creation of the system for region management. The most important and outstanding results of implementing the system in 2003 are described.

Введение

В данной работе рассматриваются проблемы взаимодействия органов власти и населения при решении социальных проблем; рассматриваются цели и задачи создания первой очереди системы управления с применением мультиагентных технологий в Самарской области, дается анализ соответствия целей и результатов разработки целям Федеральной целевой программы «Электронная Россия (2002-2010 гг.)», представляются наиболее значимые результаты внедрения системы управления регионом в 2003 году.

1 Проблемы эффективного взаимодействия населения и органов власти

В любой современной стране, независимо от особенностей ее государственного устройства и уровня экономического развития, государство – основной поставщик услуг населению, организациям и всем ветвям власти. Государственные учреждения являются источниками нормативных актов, лицензируют многие виды деятельности, обеспечивают социальную защиту населения и выполняют множество других общественно-необходимых функций. Кроме того, государство представляет собой и крупнейшего потребителя продукции и услуг частных и го-

сударственных организаций. Таким образом, для государственных органов существует необходимость постоянного взаимодействия с чрезвычайно обширной аудиторией.

Все чаще в модели взаимодействия государства и его граждан, построенной по принципу "человек обязан государству", нужно переходить к более современной, основанной на понимании того, что государство служит своим гражданам, является для них "сервисной организацией". Но при этом требуется найти разумный баланс между максимально возможным количеством новых сервисов и эффективностью их использования.

Однако для построения данной модели необходима кардинальная ломка сложившихся бюрократических подходов, когда исполнение формальных функций заменяет содержательный, заинтересованный подход к решению конкретных проблем каждого отдельного человека. Очевидно, что сама по себе автоматизация функций государственных структур не гарантирует достижения положительных результатов, если деятельность этих органов государственной власти не будет направлена непосредственно на улучшение жизни граждан, являющееся, в конечном счете, главным критерием эффективного управления регионом.

К сожалению, в настоящее время многие органы исполнительной власти действуют в значительной степени по бюрократическим принципам, взаимодействие их друг с другом и согласованность действий явно не достаточны, несмотря на понимание важности этой задачи и стремление к объединению усилий в обеспечении согласованного принятия решений и адресного взаимодействия с населением области. Большинство проблем, возникающих в результате, связано с отсутствием должного взаимодействия как между государственными структурами, так и между государственными структурами и населением, а также отсутствием обратной связи на решения принимаемые на верхнем уровне и на запросы, поступающие от граждан. В таблице 1 описаны основные проблемы, возникающие при названных видах взаимодействия, а также возможные пути их решения.

Таблица 1 – Описание проблем в управлении социальной сферой и возможные пути их решения

№ п.п.	Описание проблемы	Примеры жизненных ситуаций	Возможные решения
1	Проблема отношений человека и государства в социальной сфере	Женщина 75 лет, потеря мужа – инвалида войны, потеря двух сыновей, два инфаркта. Никогда не обращалась за помощью в социальную защиту. По выписке из больницы обратилась за льготной путевкой в санаторий (имеет пенсию 3 т.р. при цене путевки 15 т.р.). Получила ответ о возможности предоставления путевки через три года.	Принять во внимание потребности и возможности, предысторию обращений и оказанных услуг, тяжесть ситуации. Перераспределить имеющиеся путевки, обратившись к тем очередникам, для которых проблема не так остра или которые согласятся ждать ради улучшения качества предоставляемых услуг.
2	Проблема низкого качества и эффективности социальных сервисов для населения	В Самарской области, где 3 млн чел., аппаратов УЗИ больше, чем во всей Франции, где 50 млн. чел. В Самарской области 3 томографа стоимостью по 2 млн дол., которые используются менее, чем на 20% каждый. В области 80 крупных госпиталей, хотя в Израиле, где медицина считается лучшей в мире, их всего 15 (при 5 млн чел.)	Заменить иерархическую структуру на сетевую. Изменить принцип финансирования на принцип прихода «денег вслед за человеком». Обеспечить гибкое планирование и распределение ресурсов на основе анализа потребностей и возможностей населения.

№ п.п.	Описание проблемы	Примеры жизненных ситуаций	Возможные решения
3	Проблема де-бюрократизации социальных департаментов и учреждений	Количество чиновников продолжает расти, но качество выполнения их функций остается на прежнем уровне, функции часто дублируются разными ведомствами, определить лицо, ответственное за то или иное решение часто проблематично.	Административная реформа, включающая разделение и пересмотр функций учреждений, устранение избыточных хозяйственных и других несвойственных функций и т.п.

Указанные в таблице 1 пути решения социальных и управленческих проблем влекут за собой принятие новых управленческих решений, которые должны дать положительные результаты (см. таблицу 2).

Таблица 2 – Новые управленческие решения и ожидаемые результаты при решении сложившихся в социальной сфере проблем

№ п.п.	Описание проблемы	Новые управленческие решения	Ожидаемые результаты
1	Проблема отношений человека и государства в социальной сфере	Адресный подход. Замена льгот денежным пособием. Переход к взаимодействию с населением. Перераспределение помощи действительно нуждающимся.	Снижение социальной напряженности. Устранение нежелания входить в сложное положение людей и решать их проблемы. Учет мнения граждан. Рост знания своих прав.
2	Проблема низкого качества и эффективности социальных сервисов для населения	Государство как открытая сеть социальных сервисов. Департамент как центр принятия решений по развитию сети. Гибкое планирование и распределение ресурсов в сети	Повышение качества и эффективности сервисов. Экономия бюджетных средств. Больше услуг населению за те же деньги. Открытость власти.
3	Проблема де-бюрократизации социальных департаментов и учреждений	Разгосударствление социальной сферы. Выделение агентств. Расширение полномочий по принятию решений. Введение регламентов взаимодействия. Оценка по конечному результату	Сокращение аппарата. Уменьшение затрат на управление. Упрощение и ускорение решений. Повышение оплаты и мотивации чиновников.

Для решения проблем взаимодействия органов власти и населения, органов власти и бизнеса и органов власти между собой разработана Федеральная целевая программа «Электронная Россия (2002-2010 гг.)». Одной из основных ее целей является обеспечение прав граждан на свободный поиск, получение, передачу, производство и распространение информации. Для обеспечения этой цели в рамках Программы предусмотрено проведение большого количества мероприятий по обеспечению открытости деятельности органов государственной власти и общедоступности информационных ресурсов государства, созданию условий для эффективного взаимодействия между органами государственной власти и гражданами на основе использования информационно-коммуникационных технологий.

2 Внедрение в Самарской области первой очереди системы управления регионом

2.1 Цель создания системы управления регионом и решаемые задачи

Одним из региональных проектов ФЦП «Электронная Россия 2002-2010 гг.» стал проект «Разработка первой очереди системы управления регионом с применением мультиагентных технологий». Этот проект был подготовлен Администрацией Самарской области и в 2002 году представлен на общероссийском конкурсе, проводимом Министерством экономического развития и торговли Российской Федерации. В результате проведенного конкурса проект занял первое место, и Самарская область была определена опорной зоной реализации проекта.

Заказчиками реализации проекта выступили Министерство экономического развития и торговли Российской Федерации и Администрация Самарской области. Тендер на разработку выиграла в октябре 2003 г. научно-производственная компания «Генезис знаний» (г. Самара), специализирующаяся на внедрении программных продуктов с применением мультиагентных технологий.

Основная цель, которая преследуется при создании системы – повышение качества и эффективности управления в регионе для реализации потребностей и возможностей граждан и обеспечения устойчивого социально-экономического развития региона.

Эта цель достигается на основе разработки и применения радикально новых социальных, управленческих и информационных технологий, а, именно: индивидуального подхода к решению проблем каждого человека в условиях ограниченных ресурсов; нового подхода к управлению, заключающегося в замене принципа воздействия власти на население, принципом взаимодействия; использования мультиагентных технологий; перехода от систем, основанных на данных, к системам, базирующимся на знаниях.

Разработка первой очереди системы управления регионом с применением мультиагентных технологий способствует решению следующих задач Федеральной целевой программы «Электронная Россия»:

- совершенствование деятельности органов государственной власти и местного самоуправления;
- совершенствование взаимодействия органов государственной власти и органов местного самоуправления с хозяйствующими субъектами;
- повышение информационной открытости органов государственной власти и местного самоуправления, эффективности их взаимодействия с гражданами и организациями, качества оказываемых им услуг.

Система управления регионом ориентирована на следующие целевые группы пользователей, вовлекаемые в процессы адресного взаимодействия населения области и органов исполнительной власти на различных уровнях:

- граждане Самарской области,
- представители органов исполнительной власти регионального уровня,
- сотрудники подведомственных организаций органов исполнительной власти и коммерческих структур, действующих в социальной сфере.

Для граждан Самарской области система предоставит следующие возможности:

- 1) обеспечит большую эффективность во взаимодействии населения с органами исполнительной и законодательной власти региона по вопросам предоставления социальных услуг; позволит предоставлять комплексную адресную поддержку гражданам, в индивидуальном порядке решая их проблемы;
- 2) обеспечит посредством Интернет-киосков общественного доступа к системе «единое окно» для граждан для получения жизненно важной для них информации и взаимодействия с органами власти, во многом избавляя от необходимости хождения по инстанциям и простаивания в очередях;

- 3) будет способствовать повышению качества предоставляемых населению социальных услуг, а также снижению их стоимости; обеспечит гражданам постоянное информирование о состоянии рынка социальных услуг с учетом их индивидуальных потребностей и возможностей.

Для представителей органов исполнительной власти регионального уровня система предоставит следующие возможности:

- 1) будет способствовать введению новой прогрессивной модели управления социальной сферой региона, единых критериев, стандартов и подходов к управлению, базирующихся на оценке реальных услуг, отпущенных населению, и конечных результатов деятельности учреждений; сократит распыление денежных средств и повысит эффективность использования бюджетных средств;
- 2) обеспечит поддержку административной реформы, связанной с пересмотром и уточнением функций и регламентов взаимодействия социальных департаментов для повышения качества и эффективности предоставляемых населению социальных услуг;
- 3) обеспечит единое информационное пространство в регионе, интегрируя разрозненные и часто несовместимые базы данных различных социальных учреждений;
- 4) будет способствовать устранению дублирования функций государственных учреждений в социальной сфере и, наоборот, пробелов в реализации важных для населения услуг;
- 5) будет способствовать замене льгот и компенсаций единым денежным пособием для граждан.

Для сотрудников подведомственных организаций органов власти и коммерческих структур, действующих в социальной сфере открываются следующие возможности:

- 1) повысит активность организаций, действующих в социальной сфере, за счет повышения конкуренции между подведомственными органам исполнительной власти организациями и организациями коммерческого сектора;
- 2) обеспечит возможности для мониторинга состояния населения Самарской области в социальной сфере, включая социальную защиту, здравоохранение и культуру, а также степени удовлетворенности населения качеством и эффективностью предлагаемых услуг;
- 3) обеспечит возможность выхода напрямую на человека благодаря заявленным человеком потребностям и возможностям.

2.2 Задачи и результаты, достигнутые в 2003 году

В 2003 году завершился первый этап создания системы управления регионом. Весь период разработки системы продлится до 2008 года.

Разработанная в 2003 году система управления регионом включает в себя три блока.

Мультиагентная система адресного взаимодействия населения и органов государственной власти в социальной сфере, которая работает на основе карты социального паспорта. Каждый обладатель социального паспорта получает возможность из общедоступного Интернет-киоска или при помощи Интернета, из любого места, отправить запрос на предоставление льготы или услуги, а также описать свои культурные потребности. Работа программных агентов системы позволяет гражданину в режиме реального времени получать предложения от организаций с учетом индивидуальных предпочтений и формировать из них собственный календарь событий.

Мультиагентная система управления департаментом социального блока, в которой действуют менеджеры по работе с населением и менеджеры по организации процессов управления. Роль первых – непосредственно работать с людьми и помогать гражданам реализовать свои запросы, притом, что менеджеры по работе с населением лично отвечают за решение возникающих у населения проблем и за предупреждение развития проблемных ситуаций. Роль вторых – задать правила деятельности менеджеров по работе с населением и проконтролировать эффективность их работы и работы подведомственных организаций в режиме реального времени.

Интернет-портал для интеграции ресурсов и межведомственного взаимодействия – позволяет гражданам и организациям описывать свои потребности и возможности; информирует пользователей о самых лучших путях решения их проблем и реализации возможностей. В условиях ограниченности ресурсов портал помогает государственным служащим вовлекать сферу бизнеса в решение государственных задач, предоставлять гражданам широкий спектр качественных услуг в зависимости от индивидуальных предпочтений.

Работа системы управления регионом строится на основе использования баз знаний и мультиагентных технологий. В частности, в подсистеме адресного взаимодействия на текущий момент действует база знаний законов, объединяющая более 40 законов федерального, регионального и муниципального уровней и хранящая знания о более 300 льготах и выплатах и более 100 связанных с ними социальных категориях. На основе запроса от человека, его агент осуществляет оперативный поиск по базе знаний и представляет только ту информацию, которая касается конкретного человека. Подсистема управления социальным блоком кроме законодательной базы знаний использует также и базу знаний лечения заболеваний, основанную на общемировых клинических протоколах. В настоящий момент в системе действует база знаний лечения бронхиальной астмы, которая помогает врачу правильно поставить диагноз и назначить методику лечения. Агенты в данном случае помогают сопоставлять соответствие поставленного диагноза и предложенной методики лечения заложенным в базе знаний клиническим протоколам. Портал интеграции ресурсов использует базу знаний субъектов и объектов социальной сферы и связанных с ними потребностей и возможностей. Такая база знаний помогает мгновенно найти соответствующую заявленной потребности возможность и наоборот, а также получать предложения от содержащихся в портале организаций с учетом настройки персонального агента.

Адресное взаимодействие населения и органов власти осуществляется на основе социального паспорта гражданина Самарской области, интегрирующего базы данных социальных департаментов Самарской области. Социальный паспорт является совокупностью социально-значимых сведений о человеке, определяющих его общественный, имущественный и правовой статусы, хранящихся в электронной форме в информационных ресурсах органов исполнительной власти. Доступ к информационным ресурсам осуществляется при помощи специальных ключей, хранящихся в центральной базе данных системы. Система управления регионом использует существующие информационные ресурсы департаментов социального блока, интегрируя их в одну систему. К апрелю 2004 года подключены существующие базы данных министерства труда и социального развития населения Самарской области, база данных пенсионного фонда, которые хранят информацию о 1,5 млн жителях области, база данных городской поликлиники № 15 г. Самары. Интеграция с существующими базами данных создает возможность разработки на основе данного проекта системы персонального учета населения.

Граждане обращаются к системе управления регионом при помощи общедоступных Интернет-киосков, размещенных в министерстве труда и социального развития населения Самарской области, в городской поликлинике № 15 г. Самары, в музее им. П.В. Алабина, в управлении социальной защиты населения пос. Красный Яр Самарской области. В настоящее время для жителей Самарской области выпущено 500 социальных смарт-карт, открывающих доступ к системе через указанные Интернет-киоски. При помощи подсистемы управления регионом осуществляется поддержка процессов принятия решений в органах власти. В настоящий момент установлено восемь автоматизированных рабочих мест менеджеров по работе с населением и управления департаментами, установленными в министерстве труда и социального развития Самарской области, управлении социальной защиты населения Красноярского района Самарской области, городской поликлинике № 15 г. Самары, министерстве здравоохранения Самарской области, агентстве социокультурных технологий Самарской области, министерстве культуры, молодежной политике и спорта Самарской области. Интернет-портал дает возможность гражданам и организациям описать собственные возможности и потребности и уже сего-

дня получать предложения от 415 организаций, сведения о которых хранятся в портале. По сферам организации дифференцируются следующим образом: представлено 72 организации из сферы социальной защиты, 23 – из сферы здравоохранения, 257 – из сферы культуры, 63 – из других сфер.

Таким образом, в 2003 году при создании системы управления регионом *решены следующие задачи:*

- разработана и реализована первая версия мультиагентной системы адресного взаимодействия населения и органов исполнительной власти в социальной сфере (на основе социального паспорта);
- разработано программное обеспечение автоматизированных рабочих мест “Менеджер по работе с населением” и “Менеджер по организации процессов управления” департамента социального блока;
- разработан аппаратно-программный комплекс “Интернет-киоск” и установлены Интернет-киоски для свободного доступа населения в учреждениях социальной защиты, здравоохранения и культуры (4 штуки);
- разработан открытый мультиагентный Интернет-портал, обеспечивающий интеграцию социокультурных ресурсов и организацию взаимодействия организаций социальной сферы;
- создана база знаний законов в сфере социальной защиты населения;
- установлены Интернет-сервера приложений, подключены базы данных департамента социальной защиты населения, территориального фонда обязательного медицинского страхования по Самарской области, городской поликлиники №15 г. Самары, обеспечены возможности взаимодействия указанных систем посредством сети Интернет;
- выпущена пробная партия смарт-карт социального паспорта для идентификации граждан и контроля доступа к персональной информации.

Решение указанных задач позволило достичь *следующих результатов:*

- Повышение открытости, доступности и достоверности информации по социальным льготам и выплатам, повышение правовой грамотности населения, предоставление гражданам единого доступа к открытой, полной и адресной информации о социальных законах федерального, регионального и муниципального уровней;
- Сокращение времени и упрощение механизма оформления документов, подтверждающих право на получение той или иной социальной льготы, сокращение очередей;
- Обеспечение единого информационного пространства и более эффективного обмена информацией между департаментами и подведомственными им организациями в социальной сфере;
- Обеспечение прозрачности управления и предоставления услуг гражданам в социальной сфере.

2.3 Перспективы развития системы управления регионом

В 2004 году после апробации системы предполагается доработать ее функциональные и сервисные возможности: разработать на базе типового АРМ “Менеджер по организации процессов управления” специализированных по видам деятельности версий (финансы, организация управления), интегрировать типовые АРМ “Менеджер по работе с населением” с существующими программными комплексами приема населения в социальной защите и здравоохранении, подключить новые информационные ресурсы в сфере социальной защиты населения, здравоохранения и культуры (МИАЦ, ФАРМБОКС и др.), расширить перечня доступных сервисов для человека в системе адресного взаимодействия населения и органов исполнительной власти в социальной сфере, поставить дополнительные Интернет-киоски, проводить обучение пользователей работе с системой, проработать вопросы информационной безопасности.

Основным направлением работ 2005 года является доработка и внедрение системы в выбранном муниципальном образовании Самарской области. Цель такого внедрения в отдельно выбранном муниципальном образовании с компактным проживанием жителей – провести важный эксперимент и интегрировать вокруг социального паспорта все жизненно важные социальные сервисы, чтобы обеспечить для человека наибольшие удобства, открытый, адресный и эффективный подход, отражающий реальную заботу реформируемого государства о каждом человеке. Опыт такого внедрения позволит в дальнейшем быстро доработать и усовершенствовать систему, и начать ее массовое внедрение в городах и районах области в последующие годы.

Заключение

В настоящей работе рассмотрены основные проблемы, касающиеся взаимодействия органов власти и населения при предоставлении услуг населению. Большинство проблем, возникающих в результате, связано с отсутствием должного взаимодействия как между государственными структурами, так и между государственными структурами и населением, а также отсутствием обратной связи на решения принимаемые на верхнем уровне и на запросы, поступающие от граждан. Для решения этих проблем создана Федеральная целевая программа «Электронная Россия 2002-2010 гг.» и одним из региональных мероприятий, проводимых в рамках программы и решающих часть задач программы, является проект Самарской области «Разработка первой очереди системы управления регионом с применением мультиагентных технологий». Планируемый срок реализации проекта с 2003 по 2008 год. В пилотной версии 2003 года уже решена часть задач по повышению эффективности взаимодействия органов власти и населения при получении информации и услуг. В основной части статьи описаны главные результаты, достигнутые в 2003 году, а также перспективы, связанные с созданием подсистем административных регламентов и персонального учета населения в рамках развития системы управления регионом.

Список литературы

- [1] Виттих В.А., Волхонцев Д.В., Гриценко Е.А., Светкина Г.Д., Скобелев П.О. Разработка мультиагентной системы для моделирования процессов взаимодействия населения и органов исполнительной власти в социальной сфере // Труды IV Международной конференции по проблемам управления и моделирования сложных систем (Самара, 17-24 июня 2002). – Самара: СНЦ РАН, 2002. С. 414-423.
- [2] Батищев С.В., Виттих В.А., Генералова Л.Д., Горбунова Т.Ф., Куруленко Э.А., Лахин О.И., Скобелев П.О. Мультиагентный Интернет-портал по социокультурным ресурсам Самарской области // Труды IV международной конференции по проблемам управления и моделирования сложных систем (Самара, 17-24 июня 2002). – Самара: СНЦ РАН, 2002. С. 449-458.

ОПЫТ РЕФОРМИРОВАНИЯ ЗДРАВООХРАНЕНИЯ И ПУТИ ДАЛЬНЕЙШЕГО РАЗВИТИЯ ОТРАСЛИ В УСЛОВИЯХ АДМИНИСТРАТИВНОЙ РЕФОРМЫ (НА ПРИМЕРЕ САМАРСКОЙ ОБЛАСТИ)

Г.И. Гусарова¹, Г.П. Котельников², В.В. Павлов¹, В.А. Виттих³, С.И. Кузнецов¹,
А.В. Захарова⁴, Т.В. Тяпухина⁵

¹Министерство здравоохранения Самарской области
443010, Самара, ул. Чапаевская, 181, Россия
depart@medlan.samara.ru
тел: +7 (8462) 32-21-21, факс: +7 (8462) 33-45-09

²Самарский государственный медицинский университет
443099, Самара, ул. Чапаевская, 89, Россия
info@samsmu.ru
тел: +7 (8462) 32-16-34, факс: +7 (8462) 33-29-76

³ Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru
тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

⁴Государственное предприятие «Перспектива»
443010, Самара, ул. Чапаевская, 181, Россия
depart@medlan.samara.ru
тел: +7 (8462) 33-63-33

⁵Самарский областной медицинский информационно-аналитический центр
443095, Самара, ул. Ташкентская, 159, Россия
miac@medlan.samara.ru
тел: +7 (8462) 56-17-95, факс: +7 (8462) 56-17-95

Ключевые слова: реформирование здравоохранения, реструктуризация здравоохранения, государственное управление, административная реформа, эксперимент, агентства, управленческая результативность

Abstract

This paper describes the experience of federal and regional (by the example of the Samara Region) reforming of public health services and a way of its perfection in view of administrative reform spent for the country. The estimation of opportunities of optimization of public health services system is given by differentiation of functions and introduction of parameters of productivity of management.

Введение

Здравоохранение – сложная функциональная система, нуждающаяся в новых подходах к ее управлению. Являясь особой сферой деятельности, реализующей права граждан на жизнь и здоровье, здравоохранение должно стать приоритетным направлением экономической и социальной политики государства [1]. Грамотное и эффективное использование средств и влияния отрасли способно существенно повысить пределы ее возможностей [2].

В 1972 году Всемирная организация здравоохранения признала существующую в СССР систему здравоохранения лучшей и рекомендовала ее для распространения в других странах. К сожалению, советская модель являлась крайне затратной, и в современных условиях приходится искать более эффективные экономические подходы к управлению деятельностью отрасли.

Этапы реформирования системы здравоохранения

За последние 30 лет неоднократно предпринимались попытки совершенствования отечественной системы здравоохранения, включая финансирование, планирование и стимулирование деятельности медицинских учреждений. Основными задачами реформ являлись: повышение качества медицинской помощи, эффективности труда, а также рациональное расходование финансовых и материальных ресурсов отрасли.

Первый эксперимент был проведен в 1967-1969 годах (1 этап) и в 1974-1975 годах (2 этап). В его основу было положено расширение прав руководителей учреждений в части формирования штатного расписания, нормативов нагрузки и возможностей материального стимулирования персонала. В 1982-1986 годах осуществлен очередной эксперимент, ставивший задачей интенсификацию коечного фонда. Эксперименты не достигли ожидаемых результатов по причине ограничения самостоятельности руководителей медицинских учреждений рамками фонда оплаты труда.

Следующий эксперимент, получивший название «новый хозяйственный механизм», был основан на принципах экономического планирования развития отрасли, именуемых сегодня менеджментом. Он проводился в Куйбышевской (ныне Самарской), Кемеровской и Ленинградской областях в 1988-1993 годах. Впервые был реализован переход к системе полного фондодержания и подушевого нормированного финансирования первичного звена медицинской помощи, к оплате за оказанные медицинские услуги и межучрежденческим расчетам, внедрена методика тарифообразования. За счет доходов, получаемых из различных источников, первичное звено осуществляло собственную деятельность, а также оплачивало медицинские услуги, предоставленные его пациентам. Неизрасходованные средства направлялись на материальное поощрение и социальное развитие коллективов[1].

Это позволило усилить экономическую заинтересованность и ответственность работников за конечные результаты своей работы, а также провести реорганизацию лечебной сети: создавались и развивались стационарозамещающие формы медицинской помощи, сокращался неиспользуемый коечный фонд. В систему здравоохранения стали привлекаться дополнительные финансовые средства.

В результате внедрения нового хозяйственного механизма в Самарской области удалось добиться существенных качественных и количественных изменений в системе здравоохранения, а именно:

- улучшились основные показатели здоровья населения и здравоохранения (снизились уровни госпитализации, общей, материнской, младенческой и перинатальной смертности, сократилась средняя длительность пребывания больного на койке, оптимизировался показатель обеспеченности койками, уменьшилось число аборт, увеличилось среднее количество дней работы койки в году);
- практически были ликвидированы очереди на госпитализацию;
- улучшилась преемственность в лечении больных на различных этапах медицинской помощи;
- сократилось общее число жалоб населения.

Этот эксперимент стал нестандартным и смелым шагом к ликвидации сложившегося в восьмидесятых годах дисбаланса между потребностями населения в медицинской помощи и имеющимися в системе финансовыми ресурсами. Он показал, что отечественная система здравоохранения имеет большой потенциал реформирования и может быть существенно улучшена [1].

Дальнейшему продвижению эксперимента помешала ситуация экономической нестабильности, высокой инфляции, наступившая во время экономического кризиса 1994 года. Появилась и приняла катастрофические размеры задолженность первичного звена медицинской помощи за оказанные услуги, что привело к ограничению доступности стационарной и специализированной медицинской помощи.

В этих непростых условиях начинался новый этап реформирования здравоохранения - становление и развитие медицинского страхования. Несмотря на трудности периода экономической нестабильности, именно внедрение нового хозяйственного механизма позволило в полной мере реализовать на территории Самарской области Закон РФ «О медицинском страховании граждан в Российской Федерации».

Введение обязательного медицинского страхования было своевременным и единственно правильным решением, позволившим уменьшить негативные последствия бюджетного дефицита периода экономического кризиса и сохранить здравоохранение как отрасль. Это был механизм дальнейшего реформирования здравоохранения, способствовавший повышению ответственности государства за охрану здоровья граждан. Была создана работающая система защиты прав пациентов, обеспечивающая их конституционные права на бесплатную медицинскую помощь. Продолжилось коренное преобразование системы взаимоотношений в сфере оказания медицинской помощи. В полной мере начал реализовываться принцип: «деньги следуют за пациентом».

В то же время в результате неравномерного распределения средств функции страхователя неработающего населения, возложенные на органы местного самоуправления, должным образом не выполнялись. При этом ответственность за ненадлежащее выполнение этих функций не была установлена законодателем.

Для выхода из сложившейся ситуации в Самарской области было принято решение о централизации средств на обязательное медицинское страхование неработающего населения с передачей функции страхователя Администрации области. Новый порядок оплаты медицинских услуг в системе обязательного медицинского страхования населения обеспечил надлежащее финансирование, в том числе раздельное - первичной медико-санитарной помощи и медицинских учреждений вторичного звена, в результате произошло четкое разделение функций и ответственности производителей, заказчиков и потребителей медицинских услуг.

Наиболее значимой вехой на этапе развития медицинского страхования стала утвержденная Правительством Российской Федерации в сентябре 1998 года Программа государственных гарантий обеспечения граждан Российской Федерации бесплатной медицинской помощью, устанавливающая единые подходы к планированию расходов на здравоохранение за счет средств бюджетов всех уровней и обязательного медицинского страхования.

Ее реализация в определенной степени обеспечивает «прозрачность» финансовых потоков, направляемых в здравоохранение, эффективность использования имеющихся ресурсов, содействует реструктуризации отрасли и участию субъектов Российской Федерации в решении вопросов охраны здоровья населения.

Совершенствование системы регионального здравоохранения было продолжено в 1999 году, когда по инициативе Губернатора Самарской области началась реформа управления, основанная на принципах «холистического подхода» - системности и самоорганизации. Был создан первый в регионе корпоративный территориальный медицинский центр, объединивший по функционально-технологическому принципу несколько медицинских учреждений одного из районов города Самары. Создание центра потребовало проведения внутреннего аудита и реструктуризации имеющихся ресурсов.

Основной целью центра являлось повышение структурной эффективности здравоохранения путем корпоративного использования имеющихся ресурсов, сознательного ограничения финансовой самостоятельности отдельно взятых лечебно-профилактических учреждений. Его задачами стало совместное исполнение медицинских услуг, а также их корпоративный заказ. В результате объединения заказа сформировался значительный по объему гарантированный поток больных, размещаемый на более выгодных условиях во внешних для центра стационарах, что позволило на те же средства пролечить больше пациентов.

Этот первый опыт объединения медицинских учреждений крупного района в самоорганизующийся «холон» показал, что такая система эффективна. Снизилась стоимость госпитализа-

ций, уменьшилось число «внешних госпитализаций» и длительность пребывания на больничной койке, увеличилось количество пациентов, пролеченных по стационарозамещающим технологиям. Это – конкретный пример реальной оптимизации системы здравоохранения, получивший в 2000-2003 годах распространение на территориях районов-сателлитов.

В здравоохранении области внедряются и другие новые организационные технологии: созданы и развиваются службы врачей общей практики, неотложной медицинской помощи в учреждениях первичного звена, стационарозамещающие виды медицинской помощи, организована централизованная закупка медикаментов и дорогостоящего оборудования.

Следующим шагом по реформированию стало участие Самарской области в эксперименте по финансированию дополнительных объемов медицинской помощи неработающим пенсионерам из средств Пенсионного фонда Российской Федерации.

Вместе с тем, при значительном объеме выполненной работы и финансовых вложений в здравоохранение должной отдачи не получено, осуществленные мероприятия не оказали существенного влияния на показатели здоровья населения. Отечественное здравоохранение, в том числе в Самарском регионе в настоящее время характеризуется наличием комплекса сложно решаемых проблем.

На фоне сохраняющейся тенденции снижения уровня здоровья населения и происходящих в стране демографических процессов отмечается несбалансированность государственных гарантий предоставления гражданам бесплатной медицинской помощи и их финансового обеспечения, которая фактически компенсируется за счет населения [3].

На пятом Всероссийском пироговском съезде врачей министр здравоохранения и социального развития РФ Михаил Зурабов признал, что его не устраивает сложившаяся ситуация в здравоохранении, отметив как положительный факт заинтересованность медиков в преобразованиях, которые намечены в процессе осуществления реформ.

Одними из причин трудного продвижения единой политики реформ являются общероссийская проблема несовершенства управления в целом и организационно-структурные особенности системы здравоохранения в частности, децентрализация, отсутствие координации действий различных субъектов общественного финансирования здравоохранения [4].

Подобное положение дел требует изменения подходов к реформированию здравоохранения и объектов реформ. Необходимым условием успешных изменений в отрасли является комплексность решения задач реформирования, включающая соотнесенность основных положений российского и международного законодательства, сочетание федеральных интересов и интересов субъектов Российской Федерации, восстановление вертикали управления в здравоохранении.

Нужны новые методы управления, адекватные условиям рыночной экономики, учитывающие международный опыт и его интеграцию с тем лучшим, что было накоплено в отечественном здравоохранении.

«Необходимо изменить отношения между управленцами и врачами и, в конечном счете, между обществом, с его институтами власти, и системой здравоохранения» [5].

Соответственно, достижение целей отраслевой программы реструктуризации здравоохранения, проект которой обсуждался на апрельских 2004 года парламентских слушаниях, возможно при условии последовательного сближения с федеральной программой административной реформы. Результатом административных преобразований должно стать создание управленческого аппарата, способного обеспечить системность и последовательность проводимых в стране реформ здравоохранения.

Тезис о повышении эффективности системы федеральных органов исполнительной власти нашел отражение в Указе Президента Российской Федерации «О мерах по проведению административной реформы в 2003-2004 годах». Одними из приоритетных направлений данной реформы являются исключение дублирования и организационное разделение функций и полномочий федеральных органов исполнительной власти.

Действенным способом оптимизации организационной структуры государственных органов, децентрализации выполнения функций и предоставления услуг является внедрение результативного бюджетирования. Контроль над результатами деятельности государственных структур необходим в целях выбора оптимального государственного управления, для чего требуется информация о результатах управленческих воздействий. Поэтому представляется необходимым внедрение в практику государственного управления показателей результативности.

Началом административной реформы в Самарской области явилось Постановление Губернатора Самарской области, принятое в 2002 году и предусматривающее переход на систему бюджетирования, которая основана на мониторинге индикаторов затрат, себестоимости, продуктивности, эффективности, результативности. Цели разработки системы индикаторов - повышение эффективности бюджетных расходов и результативности деятельности органов государственного управления.

Самарская область стала одной из территорий проведения эксперимента в сфере государственного управления. Во исполнение июльского 2002 года соглашения Администрации Самарской области и Минэкономразвития РФ губернатором Самарской области утверждена программа проведения эксперимента в области. В качестве пилотного органа исполнительной власти был определен департамент здравоохранения (ныне министерство).

Одной из задач эксперимента являлся функциональный анализ деятельности департамента, который позволил уточнить и рационализировать функции управления, а также разработать показатели управленческой результативности, являющиеся необходимым условием адекватного ресурсного обеспечения деятельности данного органа исполнительной власти.

В целях повышения эффективности и устойчивости управления здравоохранением предложено создать три агентства, задачами которых является обеспечение системы здравоохранения необходимыми ресурсами, оказание государственных услуг и выполнение исполнительно-распорядительных функций. В соответствии с вышесказанным агентства могут взять на себя неспецифические для органа государственной власти функции.

Суть разработанной в рамках эксперимента структуры управления отраслью заключается в организационном разделении функций и полномочий между департаментом здравоохранения и агентствами. На высшем уровне управления (департамент) принимаются стратегические решения, и осуществляется нормотворческая деятельность. На нижестоящих уровнях (агентства) осуществляется обеспечивающая деятельность, принимаются оперативные управленческие решения. Разделение задач управления по уровням повышает качество и эффективность управленческой деятельности.

Заключение

Таким образом, цель государственной политики в области здравоохранения, определенная в Концепции развития здравоохранения и медицинской науки Российской Федерации, может быть достигнута путем реструктуризации системы здравоохранения при одновременной оптимизации организационной структуры государственных органов.

Реформирование органов исполнительной власти, являющееся важнейшим направлением сегодняшней политики правительства, как показал проводимый в Самарской области эксперимент, будет способствовать реализации главных функций государства в области охраны здоровья граждан, выполнению социальных гарантий и дальнейшему продвижению реформ, проводимых в отечественном здравоохранении.

Список литературы

- [1] Галкин Р.А., Засыпкин М.Ю. Роль нового хозяйственного механизма в здравоохранении Самарской области // Проблемы социальной гигиены, здравоохранения и истории медицины, 2003 г., № 3, С.45-49.
- [2] Тишук Е.А., Щепин О.В. Современные проблемы качества медицинской помощи. // Проблемы социальной гигиены, здравоохранения и истории медицины. 2003 г., № 3, С.13-19.
- [3] Герасименко Н.Ф. Проблемы законодательного обеспечения охраны здоровья населения Российской Федерации. // Здравоохранение. 2001 г., № 11, С. 15-19.
- [4] Данишевский К.Д. Факторы, препятствующие внедрению эффективных технологий управления системой здравоохранения. Сравнительный анализ реструктуризации в Самарской и Тверской областях. // Главный врач, 2003 г., № 2, С.5-13.
- [5] Воробьев П.А. Развитие системы рационального лекарственного обеспечения населения Российской Федерации – больше вопросов, чем ответов. // Проблемы стандартизации в здравоохранении. 2001г., № 3, С.3-10.

ОПТИМИЗАЦИЯ ПОЛИТИКИ ЦЕН И НАЛОГОВ С ПОМОЩЬЮ КОМПЬЮТЕРНОЙ МОДЕЛИ РЕГИОНА (НА ПРИМЕРЕ ТОМСКОЙ ОБЛАСТИ)

В.Г. Напреенко¹, А.С. Нариньяни¹, Ю.А. Кузин², Н.Е. Яковлев²

¹Российский НИИ Искусственного Интеллекта,
103001, Москва, а/я 111,
тел: 158-94-30, факс: 155-45-30

²Центральный НИИ управления, экономики и информации Минатома России,
Россия, 127434, Москва, а/я 971, ЦНИИАтоминформ,
тел: 210-77-01, факс: 210-85-44, 777-96-80.

Ключевые слова: управление экономикой, региональная экономика, экономико-математические модели, налоги, цены, недоопределенная математика

Abstract

In the article we consider an economic-mathematical model of a modern Russian region. This model provides the analysis of economic prospects of the region, helps to find a compromise among interests of participants of various economic processes and gives opportunity to optimize price and tax policy. The model was developed with the help of subdefinite calculation technology, which sufficiently extends scope of possibilities of economic model researches.

Введение

Отход от централизованных принципов управления экономикой поставил Россию перед необходимостью освоить принципы экономического регулирования, используемые в развитых странах Запада. Основными рычагами управления в этом случае являются цены и налоги. Ценовое управление включает регулирование общей инфляции и цен монополий. Возникает задача оптимизации цен монополий, инфляции и налогов. При такой оптимизации необходимо правильно учитывать и соизмерять интересы всех сторон – потребителей и производителей, отдельных регионов и государства в целом. Обеспечить нужное решение задачи в рамках простых аналитических оценок не представляется возможным. В частности, такого рода оценки почти никогда не позволяют предсказывать неочевидные последствия управляющих действий, возникающие как результат взаимодействия большого числа экономических закономерностей. Возникает потребность строить многопараметрические вычислительные модели, отражающие все существенные показатели и взаимосвязи экономического процесса.

Анализ требований к современным исследованиям экономики, дополненный анализом математических возможностей моделирования, привел нас к разработке аппарата недоопределенных моделей (Н-моделей) [1-10], который позволяет:

- оперировать со всей областью возможных состояний экономики, а не с отдельными вариантами решений, как это происходит при традиционном подходе к моделированию,
- легко расширять, сокращать и менять состав рассматриваемых показателей и расчетных условий (уравнений и неравенств),
- использовать как избыточные, так и неполные (недостаточные для однозначного определения значений показателей) совокупности расчетных условий,
- отказаться от поиска алгоритма решения задачи,
- применять, помимо точных данных, интервальные (от... до...) оценки исходных показателей, отражая возможную неполноту и неточность доступной информации,

- находить границы условий, обеспечивающие достижение тех или иных целей,
- легко обнаруживать противоречия в модели (если таковые возникнут),
- значительно повысить (в сравнении с обычными методами) качество получаемых оценок.

Технология Н-моделей относится к перспективному направлению constraint programming (программирование в ограничениях), интенсивно разрабатываемому в мире уже более 10 лет.

Обычные (алгоритмические) модели решают задачу в точных числах на основе упорядоченной последовательности явных зависимостей. Н-модели дают решение в форме числовых интервалов, допуская использование неупорядоченной системы явных и неявных зависимостей.

Именно технология Н-моделей использована нами при разработке компьютерной модели экономики Томской области.

1 Основные достоинства и особенности разработанной модели

Модель ориентирована на рассмотрение широкого спектра экономических задач.

Решение сводится к построению интервальных прогнозов значений показателей. Ширину интервалов можно регулировать, что соответствует изменению доверительной вероятности расчета (рис. 1).

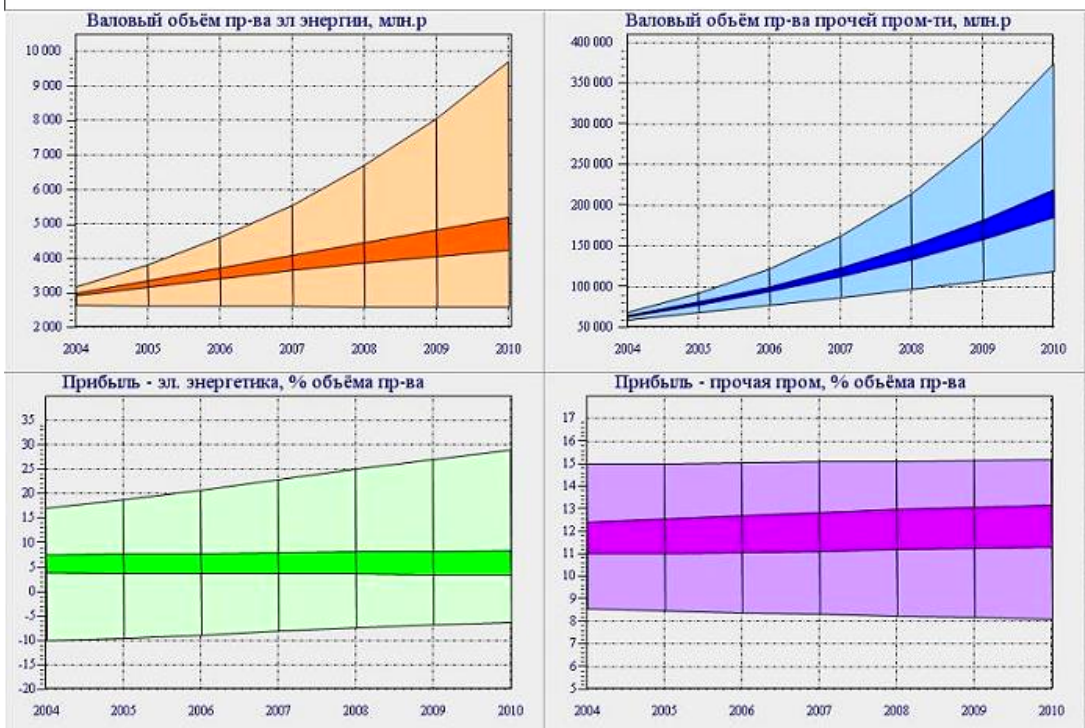


Рисунок 1 – Пример регулирования интервалов прогноза

На рис. 1 рассмотрены наложенные друг на друга две группы прогнозных интервалов – широкие интервалы, при которых область расчетных состояний экономики охватывает даже маловероятные варианты развития событий, и узкие, при которых рассматривается только наиболее вероятная ситуация.

Главной задачей считалась оптимизация политики цен и налогов в регионе. Одновременно в ходе численного эксперимента на модели обеспечено решение целого ряда смежных задач, таких как

- анализ перспектив экономики региона в зависимости от цен естественных монополий, общих темпов инфляции и величины налоговой нагрузки,
- поиск компромисса между интересами различных участников экономического процесса и выявление ограничений, позволяющих отразить эти интересы при оптимизации налогов и цен,
- индикативное планирование экономического развития региона.

Эти задачи имеют большое число вариантов, различающихся перечнями входных показателей. При обычном алгоритмическом моделировании почти для каждого такого варианта понадобилось бы создавать свою модель и находить соответствующий алгоритм вычислений. Технология Н-моделей позволила рассматривать все эти варианты в рамках единой модели, для работы которой не нужно строить алгоритм вычислений.

Чтобы использовать какой-то показатель в качестве входного показателя, достаточно указать его значение в рамках прогнозного интервала. При этом автоматически уточняются интервалы других переменных модели. В качестве примера, на рис.2 показаны результаты уточнения группы взаимосвязанных показателей – широкие интервалы на рис. 2 демонстрируют исходный прогноз, узкие – уточнение прогноза при задании пользователем части значений показателей.

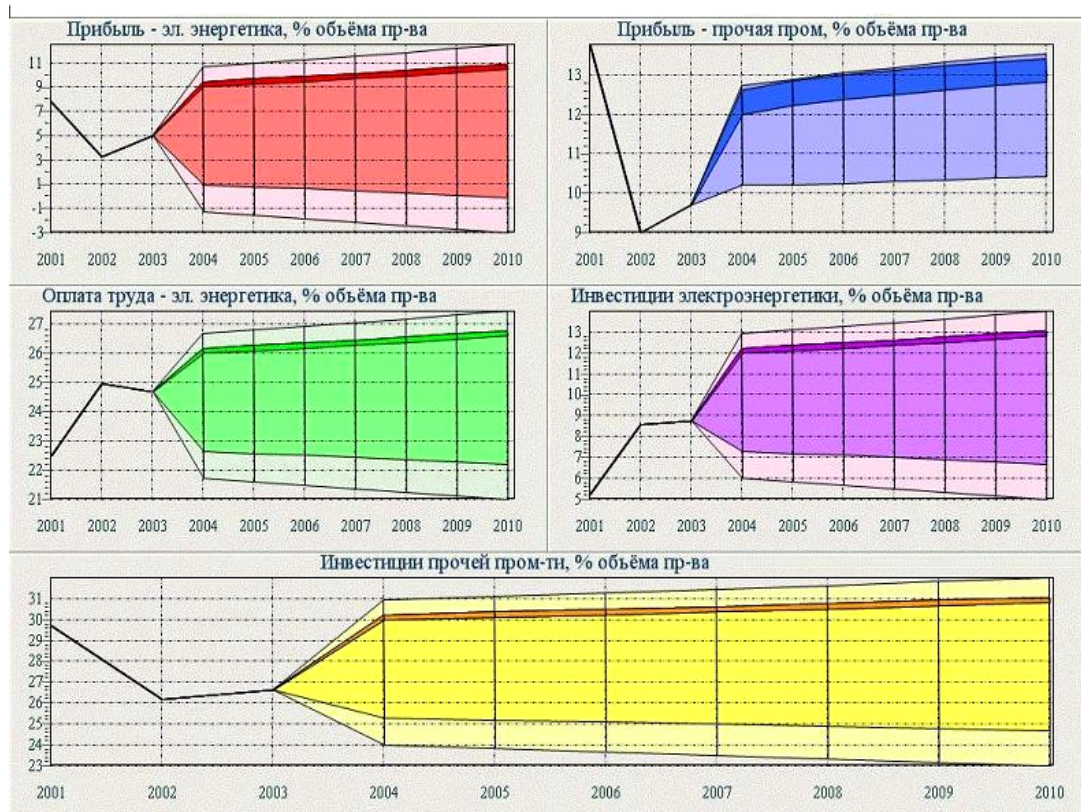


Рисунок 2 - Пример уточнения значений группы взаимосвязанных показателей

Главными показателями в рассматриваемой модели являются цены монополий, общая инфляция, параметры налогов, физические и финансовые объемы производства, оплата труда, прибыль, инвестиции, амортизация, обновление основных фондов и др.

Общероссийские и региональные изменения цен и налогов по-разному влияют на экономику региона и потому учитываются в модели отдельно.

В частности, для описания цен (тарифов) электроэнергетики в модели использованы такие показатели как средний по РФ тариф на электроэнергию для промышленности, тариф на поставки электроэнергии в регион, тариф на поставки электроэнергии из региона, тариф для региональных производителей электроэнергии, тариф для промышленных потребителей региона, коммунальный тариф региона, тариф для прочих региональных потребителей, сводный тариф по региону.

Общее число наименований экономических показателей в базовой версии модели – около 100 (без учета вспомогательных переменных). При развитии модели количество рассматриваемых показателей может быть увеличено более чем в десять раз. Наиболее сильное влияние на число переменных модели оказывает количество учитываемых секторов экономики региона.

Показатели связаны в модели большим числом соотношений (уравнений и неравенств). Важнейшими соотношениями являются:

- зависимости прибыли от цен и налогов при заданных объемах производства
- соответствие объемов производства и реализации инвестициям и ценам
- зависимости инвестиций и оплаты труда от изменений прибыли
- зависимости налоговых выплат от прибыли, оплаты труда и объемов производства
- условия натурально-вещественного и финансового баланса.

Общее число основных и вспомогательных соотношений в базовой версии модели – около 800. При развитии модели число рассматриваемых соотношений может быть увеличено до 10000 и более.

Использованные в модели соотношения позволяют отразить все существенные эффекты, возникающие при изменениях цен и налогов. Например, при росте цен естественных монополий происходит

- рост издержек производства и падение прибыли предприятий, потребляющих продукцию естественных монополий, а также рост цен на продукцию и снижение оплаты труда (например – в форме задержек зарплаты) этих предприятий, обусловленные стремления хотя бы частично компенсировать рост издержек,
- увеличение прибыли естественных монополий,
- сокращению физических объемов реализации продукции, связанное с ростом цен,
- свертывание производств, вызванное сокращением физических объемов реализации,
- изменение финансовых поступлений в бюджет.

Помимо уравнений и неравенств, описывающих моделируемые закономерности, модель позволяет использовать разного рода условия, выражающие требования к расчету. В качестве примера, на рис. 3 показаны результаты уточнения тарифов электроэнергетики и величины прибыли при задании требований к прибыли. Широкие интервалы демонстрируют исходный прогноз, узкие – уточнение прогноза при задании двух условий: прибыль электроэнергетики – не ниже минимума, близкого условию безубыточности, прибыль других отраслей – максимально возможная при соблюдении всех прочих условий прогноза.

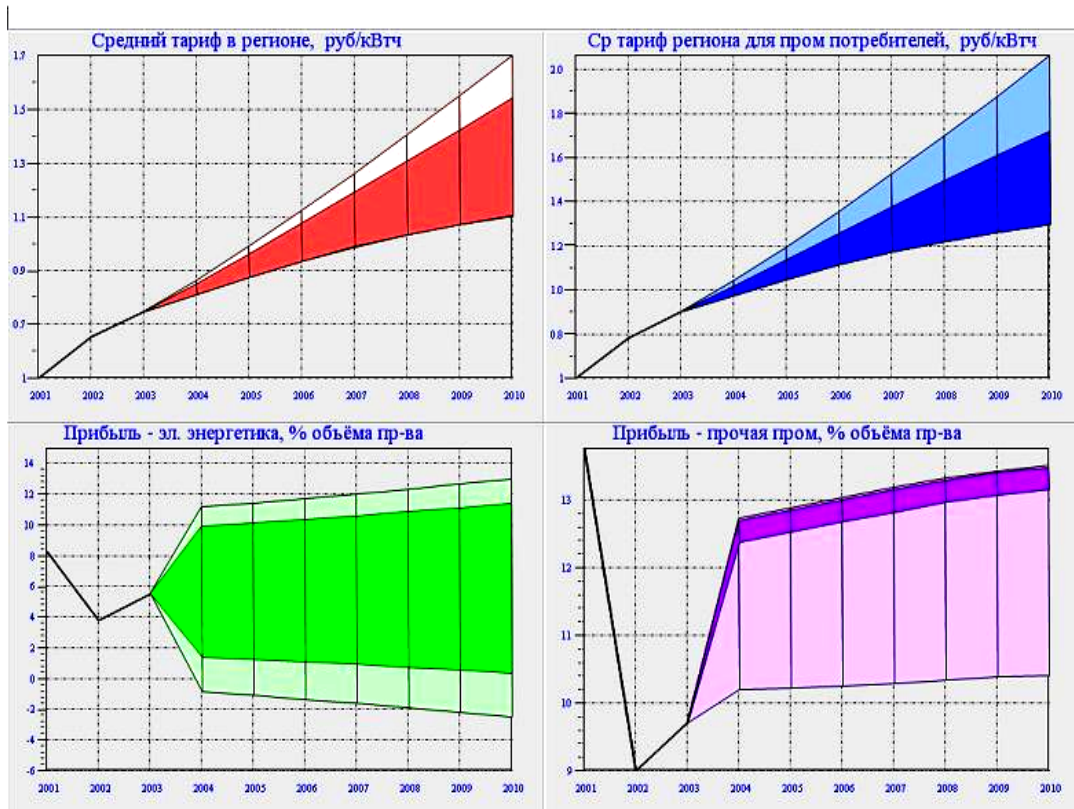


Рисунок 3.- Уточнение показателей при задании требований к прибыли

2 Работа с моделью и некоторые результаты модельных исследований

Работу с моделью обеспечивает интерфейс, выполненный в виде удобной для пользователя электронной таблицы. Основу интерфейса составила система ФинПлан [10]. С ее использованием были реализованы следующие важные преимущества:

- адекватное отражение влияния неполноты данных на результат,
- допустимость использования переопределенных зависимостей,
- одновременное рассмотрение множества вариантов решения задачи как единого решения с неточно заданными показателями и возможность использования любого показателя как в качестве задаваемого (исходного), так и в качестве выходного,
- выявление зависимости границ показателей от определенности значений различных факторов, а также автоматическое определение границ, выход за которые исключает достижение заданного результата,
- решение задач согласования данных и выявления противоречий в приближенных данных.

В итоге получена табличная модель, существенно превосходящая по своим возможностям обычные электронные таблицы (например – таблицы на базе системы Excel).

Важную роль в работе с моделью сыграли статистические данные. В начале работы по этим данным было сформировано общее представление об экономике региона, выявлена величина и структура валового регионального продукта, определена (в первом приближении) платежеспособность и финансовая устойчивость крупных и средних предприятий области, сформированы исходные оценки инвестиций в регионе и установлен вклад отраслей промышленно-

сти в экономику области. Затем с использованием статистических данных [11-14] и ряда дополнительных материалов были составлены информационные таблицы, предназначенные для настройки модели. К сожалению, многие доступные нам данные оказались низкого качества – весьма приближенными, содержащими пробелы и т.п. В этих условиях обеспечить настройку модели обычными методами весьма затруднительно. Аппарат Н-моделей легко справляется с «плохими» данными и позволяет извлекать полезную информацию даже из весьма грубых и неполных оценок. В результате работа с «некачественными» статистическими данными не вызвала у нас больших трудностей.

Основной режим работы с настроенной моделью сводится к пошаговому введению в расчет уточняющих условий (значений показателей, ограничений и др.), что ведет к сужению расчетных интервалов. Рис. 4 иллюстрирует последовательность следующих шагов

Шаг 0 – исходный прогноз, вся зона графиков,

Шаг 1- уточнены границы инфляции (более узкая зона),

Шаг 2 – задано интенсивное обновление основных фондов электроэнергетики,

Шаг 3 – введено условие максимального роста физического объема производства (внутренняя, самая узкая зона графиков).

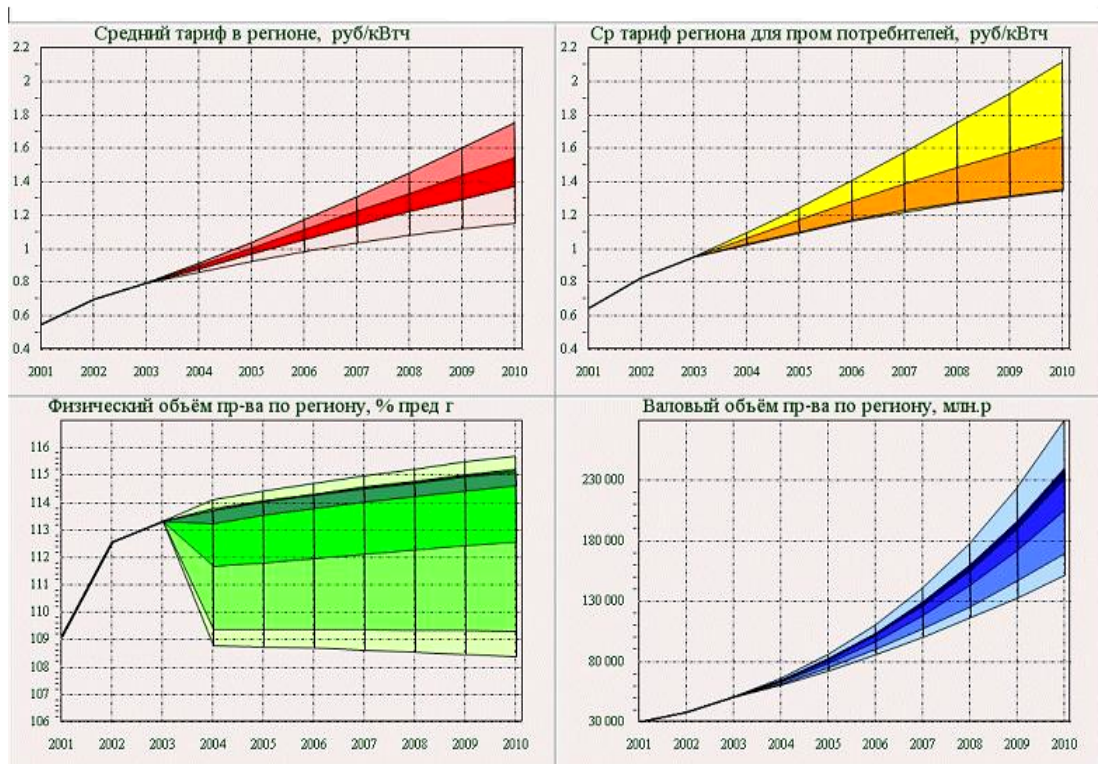


Рисунок 4 – Пошаговое уточнение расчетных показателей

В качестве важной прикладной задачи с помощью данной модели была проведена оптимизация тарифов одной из «трудных» естественных монополий Томской области – электроэнергетики. Обращение к электроэнергетике обусловлено тем обстоятельством, что последние годы в регионе проявляются две угрозы экономическому развитию. Первая угроза – низкое обновление основных фондов электроэнергетики, наблюдаемое несмотря на то, что рост энерготарифов в регионе после 2000 г опережает инфляцию. Вторая угроза – существенное ухуд-

шение экономических показателей потребляющих электроэнергию отраслей, возможное в случае, если региональные тарифы на электроэнергию будут сильно повышены.

Рассмотрим один из вариантов оптимизации.

На первом шаге задаем прогнозные значения инфляции и налоговой нагрузки, отвечающие оценкам Минэкономразвития, и выбираем заведомо широкие интервалы возможных значений тарифов.

Основными доводами в пользу роста тарифов на электроэнергию в Томской области является потребность в обновлении основных фондов электроэнергетики и несовпадение тарифов, по которым область получает и продает электроэнергию. В связи с этим в рамках интервала, полученного на первом шаге расчетов, задаем уровень обновления основных фондов электроэнергетики значениями, позволяющими в перспективе приостановить нарастание износа фондов (шаг 2) и вводим в расчет условие равенства тарифов, по которым регион получает и продает электроэнергию (шаг 3).

Выполнение шагов 2 и 3 позволяет считать, что интересы электроэнергетики нами учтены. Остается учесть интересы других участников экономического процесса. С этой целью минимизируем (насколько это возможно после сделанных шагов) тарифы на электроэнергию, сохраняя пропорции между тарифами, отраженные в статистике.

Разработанная модель показывает, что оптимальные для региона значения тарифов на электроэнергию предусматривают постепенный и ограниченный их рост, примерно на 10% в год превышающий уровень инфляции в РФ. При этом оказывается целесообразным уравнивать тарифы, по которым регион получает и продает электроэнергию. Указанные оптимальные изменения тарифов позволяют к 2007-2008гг. приостановить нарастание износа основных фондов электроэнергетики, а в последующий период дают возможность снижать износ этих фондов.

Оптимальные изменения тарифов не приводят к существенному проигрышу в показателях энергопотребляющих отраслей промышленности. Например, индекс физического объема производства этих отраслей в 2007 г снижается менее чем на 0.5% в сравнении с тем его значением, которое имело бы место при минимально возможном росте энерготарифов, совпадающем с ростом инфляции. Это снижение индекса физического объема производства можно устранить путем снижения налоговой нагрузки в регионе на 1% в год относительно средней налоговой нагрузки по России в целом.

Рост энерготарифов, существенно превышающий найденный оптимум, ведет к заметному ухудшению общих экономических показателей на всем исследованном прогнозном периоде (до 2012 г включительно). Возникающий в этом случае экономический спад не может быть устранен реально допустимым снижением налогов.

Заключение

Опыт проведенного модельного исследования может быть привлечен для построения экономико-математических моделей, решающих такие задачи как

- прогнозирование и оперативное уточнение прогнозов экономического развития регионов, отраслей, корпораций;
- долгосрочное и среднесрочное планирование деятельности предприятий, групп предприятий, отраслей промышленности;
- обоснование различных мер управления экономикой.

Список литературы

- [1] Нариньяни А.С. Недоопределенность в системах представления и обработки знаний Изв. АН СССР. Техн. кибернетика, 1986. № 5.
- [2] Дмитриев В.Е. UNICALC - интеллектуальный решатель систем алгебраических уравнений и неравенств// Искусственный интеллект-90: Тр.12 Всесоюзной конференции: - Минск, 1990. Выставка. - С. 89-91.
- [3] Narin'yani A.S., Intelligent Software Technology for the New Decade. "Communication of the ACM", vol.34, № 6, 1991.
- [4] Shvetsov, V.Kornienko, S. Preis. Interval spreadsheet for problems of financial planning. PACT'97, England, London, April 1997.
- [5] Напреенко В.Г., Нариньяни А.С., Фишкис В.М. и др. Технология компьютерного моделирования макроэкономики и бюджета Российской Федерации. М., Рос НИИ Искусственного Интеллекта, 1997.
- [6] Напреенко В.Г., Нариньяни А.С., Юртаев А.В. Недоопределенные модели – нетрадиционный подход к математическим исследованиям экономики. Информационные технологии.1999. N 4.
- [7] Напреенко В.Г., Нариньяни А.С., и др. Моделирование национальной экономики с использованием аппарата недоопределенных моделей. В сборнике "Проблемы управления и моделирования в сложных системах". Труды II Международной конференции. Самара, Самарский научный центр РАН, 2000.
- [8] Разработка недоопределенных моделей для задач экономических исследований. В сборнике трудов Седьмой национальной Конференции по Искусственному Интеллекту (23-27 октября 2000 года г. Переславль-Залесский).
- [9] Напреенко В.Г., Нариньяни А.С.. Недоопределенные балансовые модели производства. В сборнике "Проблемы управления и моделирования в сложных системах". Труды III международной конференции. Самара, Самарский научный центр РАН, 2001.
- [10] Программный пакет ФинПлан/Интегра. М. Российский НИИ искусственного интеллекта, 1997- 2003 гг.
- [11] Программа социально-экономического развития Томской области до 2005 года, утверждена на законом Томской области, №№ 384 от 28.11.2002 г./94-03 от 18.12.2002.
- [12] Томская область на пороге XXI века. Internet: <http://www.tomsk.net/>. 14.07.2003.
- [13] Официальная и деловая Россия. Справочник республик, краев, областей и округов Российской Федерации 2000/2001 – М.: Редакция издания «Официальная и деловая Россия», 2000/2001.
- [14] Статистические показатели социально-экономического положения по регионам Российской федерации. Официальные материалы Госкомстата РФ. Москва, 1997-2003.

ПРОБЛЕМЫ ОНТОЛОГИЧЕСКОЙ СПЕЦИФИКАЦИИ ОБЪЕКТОВ РЕГИОНАЛЬНОЙ ЭКОНОМИКИ

В.А. Виттих¹, Л.А. Иванова², Е.Н. Королева³,
Е.Л. Поварова², С.В. Смирнов¹ Г.Р. Хасаев²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61,
vittikh@iccs.ru

Тел. (8462) 33-26-77. Факс (8462) 33-27-70

²Министерство экономического развития, инвестиций и торговли Самарской области

³Самарская государственная экономическая академия

Ключевые слова: регион, онтологический анализ, цели развития региональной экономики, объектно-признаковые модели, отображение больших онтологий

Abstract

The examples of diverse problems which authors have collided at building of ontological specifications of objects in the regional economy are considered. Possible ways of overcoming difficulties that arise during this process are offered.

Введение

В Институте проблем управления сложными системами РАН развивается новое научное направление в области системного анализа и управления сложными системами – онтологический анализ и синтез в процессах принятия решений [1]. Одной из сфер практического применения создаваемых методов и средств является разработка онтологий региональной экономики, предназначенных для использования при формировании стратегий развития региона [2]. Начальные этапы этой работы связаны с систематизацией и формальным описанием терминов и понятий в сфере региональной экономики. Подобные задачи концептуального моделирования предметных областей могут быть решены путем построения содержательных онтологий – специально организованных компьютерных баз знаний [3].

Разработка онтологических спецификаций для актуальных сфер региональной экономики оказалась плодотворным полем сотрудничества для специалистов как в области экономики и государственного управления, так и в сфере системного анализа и обработки информации: с одной стороны обнаружилось проблемные зоны в теории и приложениях, связанных с исследованием региональных экономических систем [4], а с другой - выявляются актуальные вопросы развития методов и средств онтологического анализа сложных систем [5]. В предлагаемой статье представлены некоторые из подобных проблем, с которыми авторы столкнулись при продолжении работ по созданию онтологий региональной экономики.

1 Описания целей развития региональной экономики

Одной из животрепещущих проблем при выработке стратегий развития региона является *структуризация целей развития региональной экономики*, поскольку классификация целей по различным типологическим признакам способствует выработке более обоснованных управленческих решений.

Для региона традиционно важно связывание целей *со сферами управления* такими как: экономическая, социальная, природоохранительная, управление региональным развитием, внешнеэкономические связи (некоторые авторы добавляют еще и защиту интересов региона на

федеральном уровне [6]). А обычной практикой является импорт в региональное управление *принципов целеполагания* из традиционного стратегического менеджмента [7].

При этом в целом анализ методологических подходов к разработке региональных стратегий в России позволяет сделать вывод о пока еще преобладающем использовании концепции «регион-квазикорпорация» (т.е. когда регион, рассматриваемый как крупный субъект собственности, становится участником конкурентной борьбы на многочисленных рынках). Тем не менее, обязательным элементом общей, (главной, ключевой) цели выступают стратегические намерения, связанные с улучшением качества жизни населения регионов.

Применение формальных методов онтологического анализа [8] к структуризации целей выявило проблему *сильной связи* признаков описаний видов целей, предлагаемых в литературных источниках. На рисунке 1 представлена объектно-признаковая модель одной из подобных экспертных классификаций. Сильная связь (*пересечение* признаков, или свойств) представленных на рисунке 1 видов целей приводит к формированию в результирующей онтологической спецификации большого числа тесно связанных *общих понятий* (в рассматриваемом примере - несколько десятков), прагматическая интерпретация которых оказывается затруднительной. Фактически онтологический анализ позволяет констатировать недостаточное разделение базовых рабочих понятий в сфере структуризации целей (т.е. определенную *незрелость* этих понятий) и побуждает искать эффективные подходы к целеполаганию в региональном управлении.

Цели_развития_РЭ	Свойство_1	Свойство_2	Свойство_3	Свойство_4	Свойство_5	Свойство_6	Свойство_7	Свойство_8	Свойство_9
Цель 1	X		X	X	X	X		X	X
Цель 2		X		X			X	X	X
Цель 3	X		X	X	X	X	X	X	X
Цель 4		X		X	X	X	X	X	X
Цель 5	X	X	X	X	X			X	X
Цель 6	X	X	X	X		X		X	X
Цель 8	X	X	X	X	X	X	X	X	
Цель 9	X	X	X	X	X	X	X		X

Цели: 1 - стратегические; 2 - тактические (краткосрочные); 3 - траекторные; 4 - точечные; 5 - долгосрочные; 6 - среднесрочные; 8 - государственные; 9 - территориальных систем негосударственного типа (региональные).

Свойства: 1 - определяют общее направление развития; 2 - подразумевают достижение конкретного конечного результата; 3 - учитывают преимущественно внешние условия функционирования системы; 4 - учитывают преимущественно внутренние условия функционирования системы; 5 - разрабатываются на долгосрочную перспективу; 6 - разрабатываются на среднесрочную перспективу; 7 - разрабатываются на краткосрочную перспективу; 8 - выражают государственные интересы; 9 - выражают интересы территориальных систем негосударственного типа.

Рисунок 1 – Пример объектно-признаковой модели, описывающей классификацию целей в задачах управления региональной экономикой

2 Эмуляция содержательного описания образно определенных понятий

Проблему *отсутствия явного содержательного описания* у многих понятий, которые важны при принятии решений в сфере региональной экономики и управления, рассмотрим на примере.

Одной из фундаментальных терминологических спецификаций в экономике вообще и в региональной экономике в частности является *классификация видов экономической деятельно-*

сти (ЭД). Общероссийский классификатор видов экономической деятельности (ОКВЭД), удовлетворяющий международным требованиям, был разработан лишь в 1999 году.

ОКВЭД как нормативная терминологическая спецификация призван служить *онтологической основой* при решении следующих задач:

- классификация и кодирование видов ЭД;
- разработка нормативных правовых актов, касающихся регулирования отдельных видов ЭД;
- подготовка статистической информации для сопоставлений на международном уровне;
- определение основного и других фактически осуществляемых видов ЭД хозяйствующих субъектов;
- осуществление государственного наблюдения по видам ЭД за развитием экономических процессов;
- кодирование информации о видах ЭД в информационных системах.

Таким образом, ОКВЭД играет роль *онтологии видов экономической деятельности* и отражает эти виды путем спецификации *ресурсов* (оборудование, сырье, энергия, информационные ресурсы), которые объединяются в производственный *процесс*, имеющий целью производство продукции или оказание услуг. Другими словами, всякая ЭД характеризуется затратами на производство, процессом производства и выпуском продукции (оказанием услуг), и отличия в указанных составляющих содержательно определяет конкретный вид, или *понятие*, экономической деятельности.

Характеризуя систематизацию видов экономической деятельности в ОКВЭД, важно подчеркнуть, что в данном классификаторе использована *неявная форма указания систематизирующих признаков*, а именно: *признаки составляют набор ключевых слов предложения, которое именуется* вид ЭД. Другими словами, использование тех или иных классификационных признаков, использованных для выделения того или иного понятия, может быть идентифицировано путем, вообще говоря, слабоформализуемой обработки («дешифровки») имени ЭД. Поэтому, строго говоря, рассматриваемый классификатор, *не является содержательной онтологией*. Понятия о классах, подклассах, группах, подгруппах и видах ЭД при четком указании их иерархической взаимосвязи не имеют в ОКВЭД *характеристики (содержания)* – явного описания *состава определяющих свойств* [8, 9], который бы однозначно идентифицировал каждую разновидность экономической деятельности и формально обосновывал установленную в ОКВЭД иерархию видов ЭД.

В русле логической теории смысла *характеристика понятия* по сути дела представляет собой совокупность поименованных *измерительных процедур* (от чувственного восприятия до приборных измерений с обработкой наблюдений) [8-10]. Тогда как на примере ОКВЭД мы сталкиваемся с тем, по-видимому, распространенным в гуманитарных дисциплинах случаем, когда *описание предмета исчерпывается его именованием* на том основании, что человеку для идентификации сущностей необходимо и достаточно соотнесение исследуемых объектов с цельными, нерасчленимыми абстрактными *образами*¹, сформировавшимися в его сознании на основе усвоения личного и общественного опыта, закрепленного языковой практикой за определенными именуемыми предложениями. При этом классификационные системы, подобные ОКВЭД, - всего лишь явные представления таких взаимосвязанных образов².

¹ Понятия и образы – основные способы выражения знания. *Понятие (концепт)* описывает совокупность сущностей (объектов произвольной природы, реальных или воображаемых), обладающих *общностью* признаков (атрибутивных) структур. *Образ*, являясь подобной же структурой признаков, воспринимается как целое (например, зрительный, образ). Основное отличие образа – размытость, нечеткость признаковой структуры.

² Используя характеристику Е.К. Войшвилло [11], давшего глубокий анализ особенностей роли и формирования понятийных систем, можно сказать, что ОКВЭД в нынешнем виде «выступает лишь как средство выделения предметов» при деградации второй ипостаси понятийной системы - задачи адекватного *представления* обобщаемых в ней предметов *в научном познании*.

В технических системах оперирование образами, вообще говоря, невозможно; место образов занимают понятия-концепты, описываемые *конечными множествами признаков*. Поэтому в содержательной *онтологии, соответствующей* ОКВЭД, каждый класс, описывающий ЭД, должен обладать хотя бы *одним* явно и отдельно указанным свойством, которое позволяло бы распознавать некоторую ЭД как объект именно данного класса и наследовалось бы всеми подклассами (группами, подгруппами, видами экономической деятельности).

С учетом образной основы ОКВЭД естественным кандидатом на роль такого определяющего свойства является булева переменная, значение которой – результат выполнения некоторой виртуальной процедуры, фиксирующий принадлежность данного объекта к конкретному классу (подклассу, группе, подгруппе, виду) ЭД. Причем в качестве имени такого свойства резонно использовать *имя определяемой им экономической деятельности*.

Подобный подход к идентификации понятий получил при работе над различными онтологиями региональной экономики (в частности, он широко использовался при онтологической спецификации региональных ресурсов в части основных фондов) рабочее наименование *«принцип эмуляции образного описания класса объектов»*.

3 Блочный принцип построения и визуализации онтологий

В ходе работ над онтологическими спецификациями региональной экономики остро обостряется проблема представления пользователю *больших онтологий* (например, рассматривавшаяся в предыдущем разделе ОКВЭД-онтология насчитывает многие сотни понятий, состоящих не только в отношении обобщения, но и имеющих другие структурные связи-отношения, вызванные многочисленными примечаниями и исключениями, составляющими неотъемлемую часть этого официального классификатора). В случае графического отображения онтологических спецификаций, наиболее естественного для семантических сетей такого рода, ни масштабирование, ни фрагментарное представление выделяемых областей, ни скроллинг, ни комбинации подобных приемов вполне справиться с этой проблемой не в состоянии.

Вместе с тем, известным способом обозримой и легко контролируемой организации большого числа элементарных объектов служит их *иерархическое группирование* с введением в систему исходных объектов объекта нового типа – группы, блока, которому предписываются равные с базовыми объектами права в смысле возможности группирования. При этом действительно эффективной является *строгая иерархия* группирования объектов. Хрестоматийные примеры такой организации дают файловые менеджеры операционных систем ЭВМ.

Представляется, что введение блочной структуры способно радикально сгладить остроту представления больших онтологий. При этом блочная организация может возникать и как естественное отражение семантики предметной области (например, для ОКВЭД резонно группировать понятия каждого очередного классификационного уровня), и как результат учета психологических особенности восприятия информации (в частности, учитывать пределы кратковременной памяти человека, объем которой составляют «магические» 7 ± 2 сущности).

Блочный принцип *построения* онтологий достигается при надлежащем развитии *метаонтологии* (так или иначе присутствующей в инструментальных средствах построения онтологических спецификаций [12]) - корректным дополнением ее понятием блока, и/или может быть поддержан путем расширения *базовой модели данных* такого инструментария.

Визуализация блочной структуры со строгой иерархией на первый взгляд не представляет затруднений: каждый раз на рабочем поле отображается содержимое одного блока, и в простейшем варианте имеется возможность перехода либо в любой вложенный блок, либо - в единственный охватывающий. Однако при визуализации онтологической спецификации возникает *принципиальная трудность графической «адресации»* связей сущностей отображаемого блока (это концепты, или классы объектов, и блоки) с сущностями, размещенными в блоках, не являющихся вложенными в данный блок (или «*внешними*» сущностями).

На наш взгляд, элегантным решением этой проблемы является *стандартное «направление»* упомянутых связей к *охватывающему блоку*, образ которого так или иначе с необходимостью присутствует на рабочем поле (экране). Рисунок 2 иллюстрирует эту идею: охватывающий блок представляет на экране верхняя полоса рабочего поля (хотя помещенные на нее надписи идентифицируют онтологию в целом и отображаемый блок X), а стандартным направлением связей внутренних сущностей блока с внешними сущностями является *вертикаль*. Заметим, что это решение содержит в себе и расширение возможностей межблочной навигации, т.к. легко реализуется переход к визуализации блока, содержащего концепт, с которым имеется связь «через» охватывающий блок.

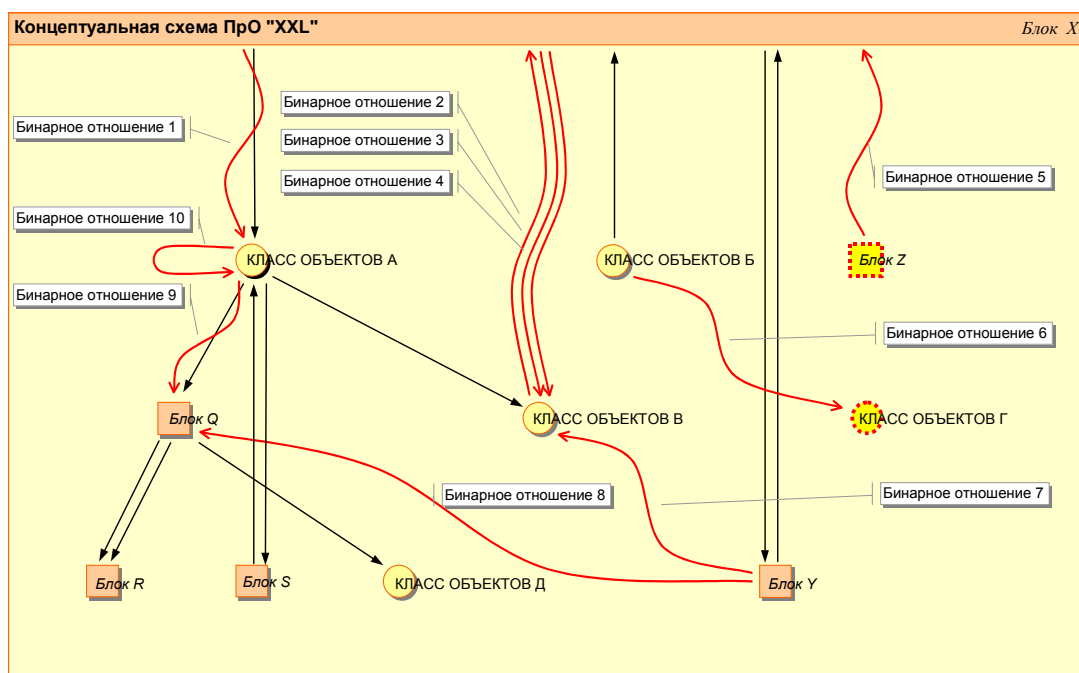


Рисунок 2 – Представление на экране содержания блока онтологической спецификации (прямые стрелки изображают отношение обобщения; Про –предметная область)

Заключение

Спецификация объектов региональной экономики с помощью методов и средств онтологического анализа способствует развитию системы знаний как в самой предметной области – в управлении региональной экономикой, так и в области компьютерных средств представления данных и знаний.

Оценивая перспективы развертывания работ по информационной поддержке процессов управления и, в частности, выработки стратегий развития региона путем построения онтологий региональной экономики следует указать три главных русла распределения усилий:

- совершенствование инструментария онтологического анализа в части представления больших онтологий, развития методов управления онтологиями и др.;
- уточнение (в результате проведенных исследований и разработок) понимания роли онтологического анализа в коллективной деятельности при управлении открытыми системами (на примере региональной экономики) с целью выработки рекомендаций по совершенствованию процессов принятия коллегиальных решений.

Список литературы

- [1] Виттих В.А. Онтологический анализ и синтез при управлении сложными открытыми системами // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, июнь 2003 г.). – Самара: СНЦ РАН, 2003. С. 50-60.
- [2] Региональная экономика как объект онтологического анализа / В.А. Виттих, Л.А. Иванова, Е.Н. Королева и др. // Известия Самарского научного центра РАН. 2003. Т. 5. № 1. С. 74-82.
- [3] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2000.
- [4] Теоретические и прикладные аспекты исследования региональной экономической системы в контексте онтологического анализа / В.А. Виттих, Л.А. Иванова, Е.Н. Королева и др. // Перспективы развития регионов в условиях глобализации: экономика, менеджмент, право: Материалы международного научного симпозиума (Самара, май 2003 г.). Ч. 1 – Самара: Изд.-во СГЭА, 2003. С. 124-128.
- [5] Онтологический анализ в задачах управления региональной экономикой / В.А. Виттих, Л.А. Иванова, Е.Н. Королева и др. // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, июнь 2003 г.). – Самара: СНЦ РАН, 2003. С. 391-401.
- [6] Гладкий Ю.Н., Чистобаев А.И. Основы региональной политики: Учебник. – СПб.: Изд-во Михайлова В.А., 1998.
- [7] Зайцев Л.Г., Соколова М.И. Стратегический менеджмент: Учебник. – М.: Юристъ, 2002.
- [8] Смирнов С.В. Онтологический анализ предметных областей моделирования // Известия Самарского научного центра РАН. 2001. Т. 3. № 1. С. 62-70.
- [9] Логический словарь: ДЕФОРТ / Под ред. А.А. Ивина, В.Н. Переверзева, В.В. Петрова. – М.: Мысль, 1994.
- [10] Рубашкин В.Ш. Представление и анализ смысла в интеллектуальных информационных системах. – М.: Наука, 1989.
- [11] Войшвилло Е.К. Понятие. – М.: Изд-во МГУ, 1967.
- [12] Смирнов С.В. Онтологическая относительность и технология компьютерного моделирования сложных систем // Известия Самарского научного центра РАН. 2000. Т. 2. № 1. С. 66-71.

ИСПОЛЬЗОВАНИЕ ОНТОЛОГИЙ ДЛЯ ПОСТРОЕНИЯ ИННОВАЦИОННЫХ ЦЕПОЧЕК В СИСТЕМЕ ПОДДЕРЖКИ ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ В РЕГИОНЕ¹

Ю.А. Загорулько¹, С.В. Булгаков²

¹Российский НИИ Искусственного интеллекта
630090, Новосибирск, пр. Лаврентьева, 6
zagor@iis.nsk.su
тел: +7 (3832) 32-83-59, факс: +7 (3832) 32-83-59

²Институт систем информатики СО РАН
630090, Новосибирск, пр. Лаврентьева, 6
bulgakov@iis.nsk.su
тел: +7 (3832) 34-29-91

Ключевые слова: Интернет-портал, информационные ресурсы, научные разработки, новые технологии и продукты, инновации, инновационные цепочки, онтология, интеллектуализация

Abstract

This paper proposes the approach to development of a program engine which provides an interactive construction of innovation chains on the Internet-portal which is intended for information support of innovation infrastructure of Siberian region. This engine is based on extensive use of system of ontologies of problem and subject domains. It is realized via a subsystem of integration of knowledge and data. This subsystem is included in the Internet-portal and provides integration of information resources concerned with innovations and intelligent information retrieval in these resources.

Введение

В новом тысячелетии только страны, восприимчивые к инновациям, смогут выстоять в конкурентной борьбе на глобальном рынке товаров и услуг. Поэтому переход к инновационной экономике очень актуален для России, имеющей большой уклон в сырьевые отрасли. Большую помощь в этом могла бы оказать российская наука, которая в лице Российской академии наук накопила огромный опыт инновационной деятельности еще в условиях планового хозяйства. В те годы внедренческие цепочки “наука-промышленность” формировались на основе административных и научно-организационных мероприятий, которые связывали академическую науку с отраслевой, с министерскими КБ и заводами в условиях стабильного бюджетного и хозрасчетного финансирования. Но с переходом к рыночной экономике организационные связи в цепочках наука - производство оказались разорванными и формируются в условиях рыночной стихии. Кроме того, после 90-го года российская экономика переживает системный кризис, в результате которого упала востребованность науки, почти в 10 раз снизилось финансирование научно-технической сферы и как следствие этого – приостановилось развитие и внедрение новых технологий. В итоге удельный вес наукоемкой продукции в российском экспорте не превышает 1,5%, что в 20 раз ниже, чем в среднем по Европе [1].

Несмотря на такие негативные тенденции, в Сибирском отделении РАН и в сибирском федеральном округе в целом за последнее время появилось много научных и научно-технических разработок, имеющих большой инновационный потенциал. В связи с этим в СО

¹Работа выполняется при финансовой поддержке Президиума СО РАН (Междисциплинарный интеграционный проект № 111).

РАН предпринимаются усилия по восстановлению и расширению сохранившихся инновационных цепочек, а также по созданию новых организационных и информационных инфраструктур взаимодействия науки и возрождающихся промышленных предприятий. Этой цели, в частности, служит проект “Сетевая инновационная инфраструктура СО РАН”, который выполняется рядом институтов СО РАН при финансовой поддержке Президиума СО РАН.

Создаваемая в рамках этого проекта система информационной поддержки инновационной инфраструктуры СО РАН должна обеспечивать широкое и активное распространение информации о взаимных возможностях (научном и/или производственном потенциале) и потребностях всех участников инновационной деятельности в регионе, чтобы создать условия для материализации накопленных знаний в виде новых технологий и продуктов.

Необходимость создания такой системы состоит в том, что уже имеющиеся системы сетевого доступа к Интернет-страницам и базам данных научных, промышленных и торговых организаций не обеспечивают нужного уровня интерактивности информационных связей и лишены аналитических, консультативных и регулирующих функций.

Система реализуется в виде Интернет-портала и включает в себя с одной стороны, информационную систему со средствами создания и интеграции связанных с инновациями разнородных информационных ресурсов, а с другой – развитые средства персонального участия в инновационной деятельности специалистов различного профиля (персонализируемые через Интернет “рабочие места” участника инновационной деятельности). Важным компонентом обеспечивающим интеллектуализации таких рабочих мест является механизм, поддерживающий интерактивное построение инновационных цепочек. В докладе рассматривается подход, при котором этот механизм основывается на активном использовании системы онтологий соответствующих задач и предметных областей.

1 Подсистема интерактивной поддержки построения инновационных цепочек

Данная подсистема является составной частью системы информационной поддержки инновационной инфраструктуры, поэтому ей будут доступны как все информационные ресурсы, интегрируемые в данной системе, так и подсистемы, обеспечивающие доступ к ним. Ее работа базируется на знаниях об инновациях и инновационной деятельности, об устройстве мира производства и науки, а также на знаниях о конкретных предметных областях, попадающих в сферу инновационной активности. Все эти знания представлены в нескольких онтологиях, имеющих самостоятельное значение, но связанных через общие понятия.

1.1 Система знаний

Система знаний подсистемы объединяет следующие онтологии: 1) онтологию инноваций и инновационной деятельности, 2) онтологию мира производства 4) онтологию науки и научного знания, а также 4) онтологии предметных областей, интересных с точки зрения инноваций и представленных в системе своими информационными ресурсами.

Онтология инноваций и инновационной деятельности строится на основе результатов исследований, представленных в [2, 3]. Она включает следующие группы понятий: инноваторы (предприятия, научные организации и т.п.), инвесторы, этапы инновационного процесса (фундаментальные исследования, НИР, ОКР, маркетинг, производство и др), объекты и субъекты инноваций и т.п. Эти понятия связаны такими отношениями, как “результат этапа инновационного процесса”, “порядок следования этапов”, “исходные данные для этапа”, “участник этапа” и т.п.

Онтология мира производства включает иерархию производственных отраслей и типов производств и предприятий, а также такие базовые классы понятий, как производитель, поставщик, ресурсы (сырье, комплектующие, трудовые ресурсы).

Онтология науки и научного знания описывает понятия, относящиеся к организации научной деятельности, а также метапонятия, задающие структуры для описания предметной об-

ласти любой конкретной научной дисциплины. К таким метапонятиям относятся: раздел науки, метод, объект и результат исследований.

Онтологии предметных областей включают понятия, характерные для конкретной предметной области, и связи между ними. При этом они могут опираться на онтологию науки и научного знания и/или онтологию мира производства. Примером таких областей могут быть: программирование, геофизика, приборостроение, металлургия, самолетостроение.

1.2 Механизм построения инновационных цепочек

Построение инновационной цепочки выполняется по автоматически генерируемому сценарию, структура которого определяется структурой инновационной цепочки, заданной в онтологии инновационной деятельности, и видом инновационного запроса.

Вид инновационного запроса зависит от типа пользователя (ученый, производитель, инвестор и т.д.) и может состоять, например, в продвижении какой-либо научной идеи или организации производства нового прибора или материала.

Инновационный запрос порождает на каждом этапе его обслуживания (продвижения по инновационной цепочке) определенную информационную потребность. Эта потребность представляется в виде поискового запроса в терминах конкретной предметной области, соответствующей этапу инновационного процесса и типу продукта инновации.

Затем этот запрос направляется к подсистеме интеграции информационных ресурсов портала, которая преобразует его в ряд запросов к конкретным источникам данных. Результаты обработки запроса «подцепляются» к информационному каркасу цепочки.

Если пользователь решает, что данный этап достаточно обеспечен информацией, он может перейти на следующий этап построения цепочки. Так как может быть получено несколько альтернативных вариантов реализации текущего этапа цепочки, то пользователь может выбрать те варианты, которые он считает наиболее предпочтительными (подсистема позволяет сохранять множество вариантов и обеспечивает возможность работы со всеми ними одновременно). Дальнейшее продвижение по цепочке будет осуществляться только с учетом полученной в рамках этих вариантов информации.

Как было сказано выше, свою инновационную потребность пользователь задает в понятиях конкретной онтологии (выбранной самим пользователем или системой). Если выясняется, что адекватных понятий в этой онтологии нет, то пользователю будет предложено описать новое понятие и указать его место в уже имеющейся системе понятий. При этом ему, возможно, придется достроить недостающую цепочку понятий от наиболее близкого по смыслу понятия, имеющегося в онтологии, к новому. В случае успешного выполнения указанной процедуры, будет выполнена достройка онтологии.

Если на портале нет подходящих онтологий, то это означает, что пользователь не сможет реализовать свою инновационную потребность на нашем портале, так как портал не располагает для этого необходимыми знаниями и информационными ресурсами. На основе анализа таких случаев будут приниматься решения о подключении к portalу новых информационных ресурсов и онтологий.

2 Подсистема интеграции знаний и данных

Рассмотрим структуру и функционирование подсистемы интеграции знаний и данных, которая обслуживает информационные потребности подсистемы интерактивной поддержки построения инновационных цепочек.

Данная подсистема включает набор модулей и компонентов, обеспечивающих интеграцию информационных источников и проведение содержательного поиска на основе онтологии. Подсистема основана на мульти-агентном подходе, что позволяет динамически подключать новые источники данных и обеспечить одновременную работу с множеством запросов к различным информационным источникам.

Подсистема интеграции ориентирована на работу со структурированными источниками данных (СИД), для которых может быть построена схема данных, то есть описаны типы данных и связи между ними, и существует формальный способ получения отдельных элементов данных. Примерами СИД можно считать различные базы данных (например, реляционные и объектные), а также слабоструктурированные ресурсы, описанные в форматах XML, RDF, OWL, DAML+OIL [4] и др.

Подобный подход к интеграции знаний и данных предложен в [5].

2.1 Интеграция источников данных на основе онтологии

Интеграция доступных СИД происходит в соответствии с используемой онтологией, что позволяет строить поисковые запросы, абстрагируясь от способа и формата хранения данных каждого СИД. Для этого необходимо решение следующих задач:

- Сопоставление схемы каждого СИД с элементами используемой онтологии.
- Согласование схем данных источников с учетом различия форматов представления данных, принятых в них.

По схеме данных каждого СИД строится модель, которая подобно онтологии представляется в виде сети понятий и связей, соответствующих элементам схемы данных. Понятия модели характеризуются атрибутами, соответствующими атрибутам конкретных типов хранимых данных. Каждому атрибуту модели соответствует домен, который кроме множества допустимых значений этого атрибута определяет еще и их формат записи, принятый в данном СИД.

После построения модели источника данных настройщик-эксперт может модифицировать ее, удаляя часть элементов или сужая домены. Это дает возможность упростить работу с источником в случае, когда необходимо использовать только часть содержащейся в нем информации.

Полученная модель связывается с онтологией: при этом для всех элементов модели ищутся соответствующие им элементы онтологии. Так как не всегда возможно установить прямое соответствие между элементами онтологии и модели СИД, то на элементы модели могут накладываться дополнительные ограничения, в частности, ограничения на множество значений некоторых атрибутов. После сопоставления атрибутов из онтологии элементам схемы данных необходимо выполнить сопоставление связанных с ними доменов. Для этого строится формула преобразования значений. Формула исполняет роль транслятора между форматами онтологии и источника данных и элиминирует различия в формах написания значений, языках, единицах измерения и т.п.

Далее, все модели источников данных объединяются в общую модель данных, отражающую информационное наполнение всех доступных источников.

2.2 Обработка запросов

После построения общей схемы данных и сопоставления ее элементов с элементами онтологии поисковый запрос может быть задан в терминах выбранной ПО. В запросе указываются тип (или набор типов) искомой информации и ограничения, которым должны удовлетворять найденные данные.

Тип искомой информации задается указанием класса, являющегося элементом онтологии, и требуемых атрибутов этого класса (возможно всех).

Ограничения могут быть заданы в виде поискового шаблона определенного вида (то есть, с фиксированными допустимыми значениями атрибутов) и/или логическими выражениями над значениями атрибутов понятий ПО. Между ограничениями в запросе могут быть установлены логические связи, позволяющие определять конъюнкцию и дизъюнкцию условий. Ограничения могут быть заданы на произвольные элементы онтологии, частично или полностью определяя их допустимые значения. Элементы, на которые наложены ограничения, будем считать *частично заданными* или просто *заданными*.

В общем случае поисковый запрос представляет собой фрагмент онтологии, включающий искомые и заданные элементы. Между искомыми и заданными элементами запроса могут быть определены связи, то есть цепочки классов и отношений, связывающие эти элементы. Явное описание таких связей позволяет четко определить требуемые отношения между искомыми и заданными элементами.

Если такие связи не заданы в запросе явно, то они могут быть достроены автоматически в соответствии с онтологией. В этом случае запрос считается *неточным*, и он уточняется (расширяется) всеми возможными цепочками связей.

Функционирование поисковой системы при обработке запроса пользователя происходит по следующей схеме. Запрос пользователя, построенный в терминах онтологии, отображается на общую схему данных, после чего порождается агент запроса, управляющий процессом сбора информации. Для доступа к информации в конкретных источниках данных используются специализированные поисковые агенты, ориентированные на работу с этими источниками. На основании найденной информации и знаний, содержащихся в онтологии, строится ассоциативная сеть, являющаяся результатом запроса.

В процессе поиска рассматриваются все допустимые с точки зрения запроса пути от вершин, для которых заданы ограничения, до искомых вершин в построенном по запросу фрагменте общей схемы данных. Сначала выбирается информация, соответствующая заданным и частично заданным элементам запроса. Далее рассматривается следующий элемент схемы данных и связь с ним в рамках выбранного пути до искомой вершины фрагмента сети, представляющего запрос. Для каждого такого перехода по звеньям пути строится элементарный подзапрос. Он передается агенту, настроенному на работу с источником, в котором содержится информация, соответствующая новой вершине. Эта информация выбирается агентом с учетом данных, найденных для исходной вершины. Порождение подзапросов и выборка информации происходит до тех пор, пока не будет достигнута вершина, помеченная в запросе как искомая, либо результат очередной выборки данных не окажется пустым.

Сбором результатов управляет поисковый агент. В процессе поиска строится сеть связанных между собой результатов, сходная со структурой самого запроса. Используя специальный интерфейс визуализации результатов поиска, пользователь может осуществлять навигацию по визуальному представлению этой сети. Подобная визуализация упрощает процедуру выбора наиболее важной информации пользователем за счет явного визуального представления связей между найденными данными. Кроме того, упрощается процедура задания уточняющих или дополняющих поисковых запросов.

3 Пример

Пусть нашим порталом хочет воспользоваться коллектив разработчиков новых электроизмерительных приборов. Предположим, что у него есть опытный образец нового прибора и желание запустить его в серийное производство.

В нашей системе онтологий такой коллектив будет выступать инноватором, а точнее разработчиком. С точки зрения инновационного процесса, он уже прошел стадии фундаментальных исследований, НИР, ОКР и в качестве текущего результата инновационного процесса имеет опытный образец.

На первом шаге инноватор должен позиционировать свою разработку, т.е. выбрать нужную онтологию предметной области и указать в ней наиболее близкое понятие, в объем которого может быть включен объект инновации.

На следующем шаге инноватор переходит к этапу маркетинга и определяет круг партнеров, которые могли бы провести исследование рынка нового прибора. При этом генерируется запрос к подсистеме интеграции информационных ресурсов на поиск организаций, которые могут выполнить маркетинг для данного класса приборов.

Следующий шаг связан с выбором инвесторов. Ими могут выступать банки, администрации, правительства, фонды и т.п. Такие организации могут быть найдены также по запросу к подсистеме интеграции данных.

Последний шаг связан с организацией производства нового прибора. Необходимо выбрать базовое предприятие, на котором будет выпускаться данный прибор, а также предприятия-поставщики комплектующих и сырья для производства нового прибора.

Информационное обеспечение этого шага – наиболее сложная задача, для решения которой может потребоваться многоступенчатая генерация и обработка информационных запросов. Так как прибор может иметь довольно сложную структуру и состоять из многих комплектующих, которые могут не только браться, грубо говоря, со склада, но и могут потребовать организации нового производства (см. рисунок 1). Для таких производств потребуются свои инновационные цепочки и т.д. Правда, как правило, такие цепочки будет гораздо проще основной, так как в них будут отсутствовать этапы маркетинга и поиска инвесторов.

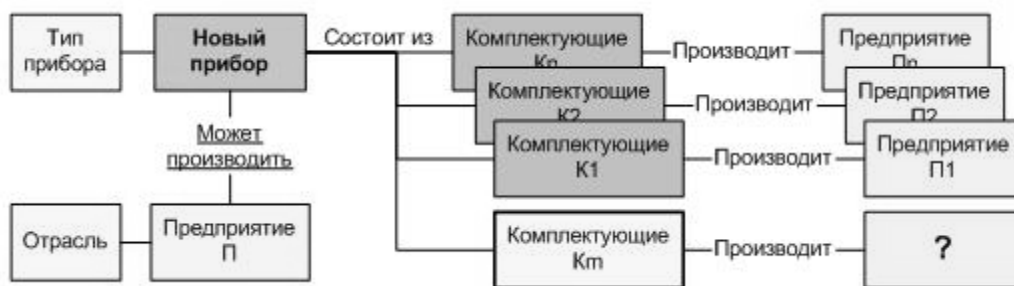


Рисунок 1 – Фрагмент схемы производства нового прибора.

Заключение

В докладе предложен механизм, поддерживающий интерактивное построение инновационных цепочек в сетевой системе информационной поддержки инновационной инфраструктуры региона, доступной пользователям в виде специализированного Интернет-портала. Этот механизм основан на активном использовании системы онтологий проблемных и предметных областей. Для его реализации используется входящая в Интернет-портал подсистема интеграции знаний и данных, с помощью которой обеспечивается интеграция связанных с инновационными разнородных информационных ресурсов и содержательный поиск в них.

В ближайшее время планируется расширение системы знаний, в частности, подключение онтологий новых предметных областей и соответствующих им информационных ресурсов.

Список литературы

- [1] Автономов В.П. Национальные инновационные системы и место территорий с высокой концентрацией научно-технического потенциала // Материалы проекта FinRus 9804 Европейского Союза, Инновационные центры и наукограды. – М., 2001.
- [2] Унтура Г.А. Модификация научной деятельности в новом информационном пространстве (создание и тиражирование научных результатов с помощью сетевой инновационной инфраструктуры). http://sinin.nsc.ru/report_SPB1.html
- [3] Зверев В.С., Унтура Г.А. Онтология инновационной деятельности. <http://sinin.nsc.ru/docs/Ontologia.doc>
- [4] <http://www.w3c.org/>
- [5] Смирнов А.В., Левашова Т.В., Пашкин М.П., Шилев Н.Г. Онтолого-ориентированный многоагентный подход к построению систем интеграции знаний из распределённых источников // Информационные технологии и вычислительные системы. 2002. № 1. С. 62-82.

РАЦИОНАЛИЗАЦИЯ ФУНКЦИЙ И ОРГАНИЗАЦИОННОЙ СТРУКТУРЫ ОРГАНОВ ГОСУДАРСТВЕННОЙ ВЛАСТИ В РАМКАХ ЭКСПЕРИМЕНТА В СФЕРЕ ГОСУДАРСТВЕННОГО УПРАВЛЕНИЯ В САМАРСКОЙ ОБЛАСТИ (НА ПРИМЕРЕ ДЕПАРТАМЕНТА ЗДРАВООХРАНЕНИЯ)

В.А.Виттих¹, Г.И.Гусарова², В.В.Павлов², А.В.Захарова³, Т.В.Тяпухина⁴,
Н.В.Дилигенский¹, С.В.Кузнецов¹

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
cscmp@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²Министерство здравоохранения Самарской области
443010, Самара, ул. Чапаевская, 181, Россия
depart@medlan.samara.ru

тел: +7 (8462) 32-21-21, факс: +7 (8462) 33-45-09

³Государственное предприятие «Перспектива»
443010, Самара, ул. Чапаевская, 181, Россия
depart@medlan.samara.ru

тел: +7 (8462) 33-63-33

⁴Самарский областной медицинский информационно-аналитический центр
443095, Самара, ул. Ташкентская, 159, Россия
miac@medlan.samara.ru

тел: +7 (8462) 56-17-95, факс: +7 (8462) 56-17-95

Ключевые слова: административная реформа, государственное управление, функциональный анализ, эксперимент, агентства, аутсорсинг

Abstract

This paper describes the functional analysis of body of the government within the framework of experiment spent for the Samara Region in sphere of government reforming. The way of optimizing the interaction between departments by delegation non-specific functions of body of the government to agencies is proposed.

Для обеспечения исполнения стоящих перед государством задач, сформулированных в Посланиях Президента страны Федеральному Собранию, требований Конституции РФ и потребностей современного российского общества наряду с рационально построенным законодательством необходимы эффективные технологии государственного управления и результативные механизмы реализации законов и политических решений [1].

Одна из ключевых проблем российского государственного управления заключается в дублировании функций органов исполнительной власти одного уровня и пересечении функций органов исполнительной власти разного уровня, осуществлении функций, непосредственно не отражающих властные prerogatives государства.

В работе [2] отмечена неэффективность системы делегирования ответственности «сверху вниз» и горизонтальных взаимодействий на нижних уровнях управления.

Налицо неконкретность, избыточность, отсутствие единообразия в определении функций, механизмов их реализации и координации действий. Происходит подмена развивающихся способов рыночного регулирования экономических процессов, а также системы саморегулирования таких процессов административным регулированием.

Наряду с этим недостаточно государственных функций в тех сферах общественной жизни, где существует конституционная ответственность исполнительной власти. В то же время сочетание в одном органе исполнительной власти разных видов функций приводит к принятию узковедомственных решений.

Поэтому рационализация функций органов государственной власти - одна из основных задач проводимой в стране Административной реформы, целью которой является придание системного практического характера продвижению эффективных технологий государственного управления, создание единого управленческого пространства [1].

Во исполнение Указов Президента Российской Федерации от 19 ноября 2002 г. «О федеральной программе реформирования государственной службы РФ» и от 23 июля 2003г. № 824 «О мерах по проведению административной реформы в 2003-2004 годах» и Соглашения Администрации Самарской области и Минэкономразвития РФ от 17 июля 2002 года 27 декабря 2002 года Губернатором Самарской области утверждена Программа проведения эксперимента в сфере государственного управления в Самарской области.

В рамках данной Программы, в целях апробации новых экономических механизмов, а также форм организации государственного управления пилотным департаментом был определен департамент здравоохранения.

Высшей школой экономики и Институтом проблем управления сложными системами РАН для проведения эксперимента была подготовлена Рабочая программа научно-исследовательских работ по теме: «Проведение экспериментов и осуществление пилотных проектов по применению новых подходов к организации федеральной государственной службы и обеспечению деятельности федеральных государственных служащих в Самарской области».

Цель работы – функциональные и организационные изменения структуры, методов планирования, бюджетирования, финансирования, мониторинга, оценки и контроля деятельности органа власти в зависимости от типов осуществляемых функций, выполняемых видов работ.

Поэтому одной из задач Эксперимента явился функциональный анализ деятельности департамента для уточнения его функций, их рационализации, устранения дублирования, обоснованного внесения изменений в его организационную структуру.

За основу взята типовая методика, подготовленная сотрудниками Центра проблем государственного управления Государственного университета – Высшей школы экономики и Экспертного Фонда социальных исследований «ЭЛЬФ».

Было проведено исследование предписанных департаменту и реально выполняемых им функций (в том числе методом опроса сотрудников), а также анализ функций его подразделений и учреждений, обеспечивающих деятельность департамента.

Анализ показал, что функции, в целом, характеризуются расплывчатостью формулировок, нечеткостью описаний действий, отсутствием ориентации на конечный результат и конкретных внешних потребителей.

Ответы на вопросы «что делает организация», «кто осуществляет эту деятельность», «зачем, почему, каким образом», «имеется ли правовое обеспечение функции» позволили уточнить, устранить избыточные, реально не выполняемые, неконкретные, дублирующие, не свойственные органу государственной власти, либо не определенные законодательством функции, а также дополнительно включить функции, не прописанные в положениях, но фактически осуществляемые департаментом [3].

Были выделены неспецифические функции, выполнение которых может быть делегировано (аутсорсинг функций).

Кроме того, традиционно используемые формулировки функций не отвечали требованиям разработанного в рамках федеральной программы «Реформирование государственной службы Российской Федерации (2003-2005 годы)» стандарта описания функций государственных органов исполнительной власти и их подразделений (отсутствовала их структуризация и классификация по целевой ориентации и типам выполняемых действий).

В соответствии с разработанным стандартом и алгоритмом анализа существующие функции были всесторонне проанализированы и структурированы по базовым определяющим признакам.

Формулировки функций коренным образом переработаны согласно установленным правилам описания функций, и им придан конструктивный характер, ориентированный на конечный результат деятельности.

Выявлены следующие базовые типы функций департамента:

1. политические функции, определяющие основные перспективные направления действий и пути их реализации;
2. международные функции, характеризующие внешние связи и отношения;
3. нормативно-регулятивные функции, устанавливающие нормативно-правовую основу деятельности;
4. исполнительно-распорядительные функции, описывающие процедуры принятия конкретных решений по отдельным объектам регулирования;
5. контрольно-надзорные функции, обеспечивающие проверку соответствия декларируемой и осуществляемой деятельности установленным законодательным и иным нормативным, правовым требованиям;
6. управления государственным имуществом, реализующие права государства, как собственника;
7. оказания услуг, выполняющие действия в интересах стороннего пользователя;
8. обеспечивающие функции, поддерживающие реализацию властных полномочий.

В процессе структуризации и классификации функций, как правило, осуществлялось деагрегирование функций и проводилась практически полная перефразировка в соответствии с их реальным содержанием. При этом одна существующая формулировка заменялась несколькими новыми, и количество стандартизированных функций существенно возросло.

Проведенный функциональный анализ позволил разработать систему показателей результативности департамента, являющихся необходимым условием адекватного обеспечения его деятельности ресурсами.

Положение о департаменте было приведено в соответствие с новыми формулировками его функций, разработанных согласно указанного выше стандарта их описания, и устранены недостатки действующего положения.

Следует отметить также, что уточнение и рационализация функций департамента, исключение их дублирования, описание процессов их выполнения нашли свое практическое применение при подготовке Положения о Министерстве здравоохранения Самарской области и Регламента его деятельности.

Существующая система управления здравоохранением Самарской области опирается на классическую линейно-функциональную организационную структуру, представленную на рисунке 1. Основой такой системы управления являются линейные подразделения. При этом деятельность линейных подразделений является функционально-ориентированной.

Линейные подразделения структуры управления выполняют четко очерченные специализированные функции оперативного руководства: организацию медицинской помощи взрослому населению, детям, родовспоможение, управление санаторно-курортной сферой, организацию фармацевтической работы. Кроме этих специфических функций в условиях функционирования здравоохранения как социально-экономической подсистемы общества базовыми функциями управления здравоохранением являются планирование, бюджетирование, учет и подготовка кадров, режимная и прочая деятельность.

Линейно-функциональная структура управления является наиболее распространенным типом систем управления, позволяющим сочетать четкость (определенность) действий и относительную автономность работы (самоорганизацию) деятельности линейных подразделений. Функциональная специализация подразделений позволяет эффективно выполнять необходи-

мые работы, повышать качество управленческой деятельности. При этом недостатком этой структуры является отсутствие гибкости, необходимой в условиях происходящих структурных перестроек управления социально-общественными процессами.

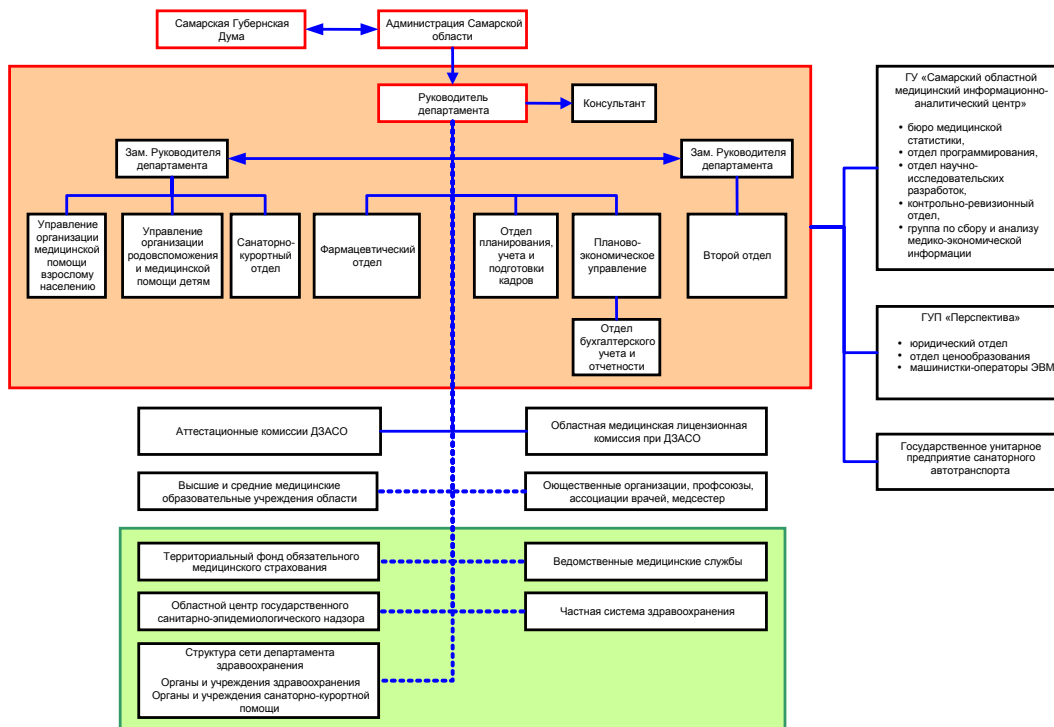


Рисунок 1 – Действующая организационная структура системы управления здравоохранением (на момент проведения эксперимента)

Были подготовлены предложения по более рациональному распределению функций между подразделениями департамента, а также по оптимизации внутриведомственного взаимодействия.

Структура системы управления здравоохранением Самарской области, разработанная в рамках федеральной программы «Реформирование государственной службы Российской Федерации (2003-2005 годы)», представлена на рисунке 2.

Оперативное управление работой департамента осуществляется непосредственно руководителем департамента и его заместителями. Управляющими органами являются три управления и четыре отдела.

В структуре управления выделены две группы подразделений, обеспечивающих суммарную эффективность системы здравоохранения. Первая выполняет организующую роль для достижения необходимого профессионального уровня и соответствия деятельности организаций и работников областной системы здравоохранения установленным требованиям. Это областной центр государственного санитарно-эпидемиологического надзора, областная медицинская лицензионная комиссия, аттестационные комиссии, высшие и средние специальные медицинские образовательные учреждения области.

Вторая группа подразделений обеспечивает эффективность организационных способов повышения качества работы системы здравоохранения: территориальный фонд обязательного медицинского здравоохранения, ведомственные медицинские службы, общественные органи-

зации (профсоюзы, ассоциации врачей, медсестер), частная система здравоохранения. Нижними уровнями организационной структуры управления являются региональные учреждения здравоохранения и санаторно-курортной помощи.

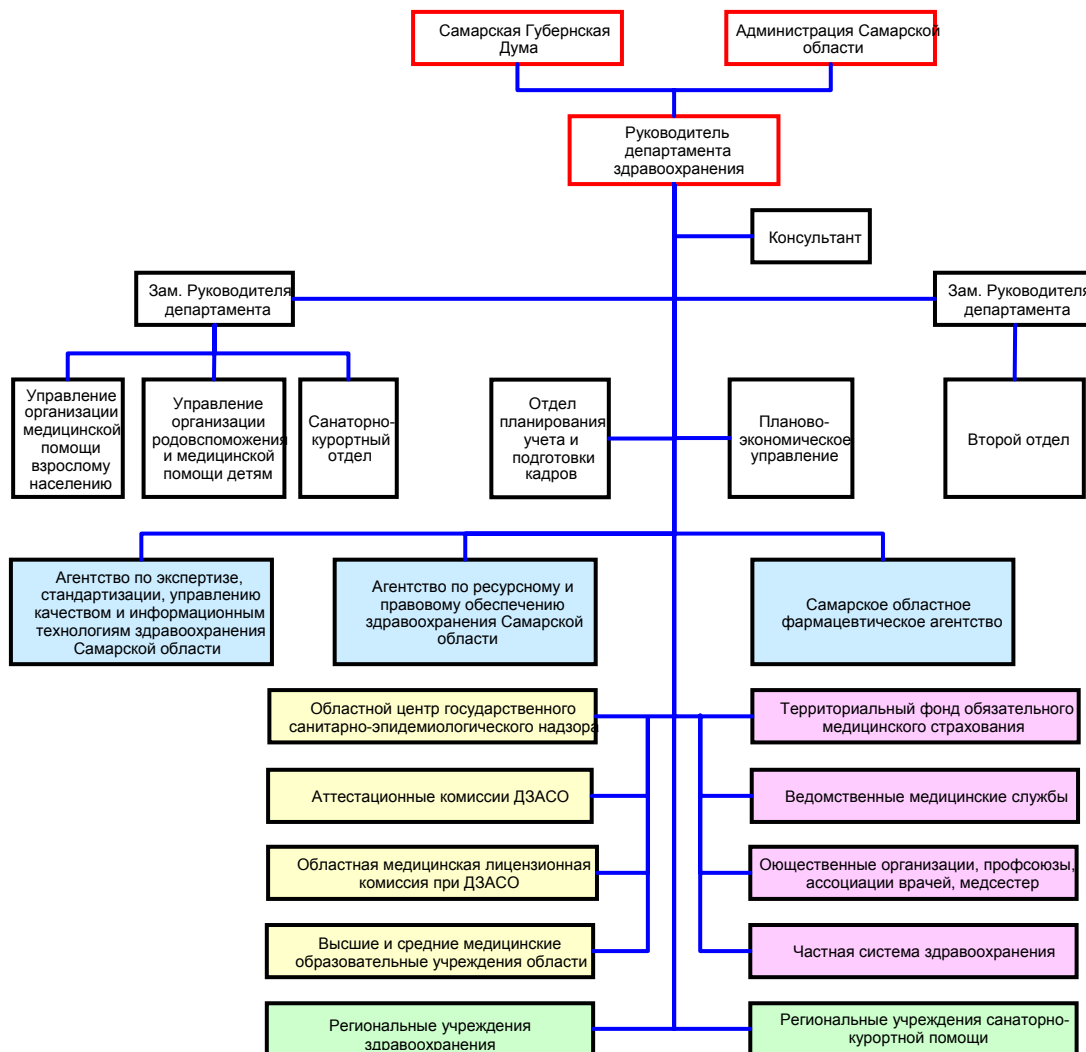


Рисунок 2 – Организационная структура системы управления здравоохранением Самарской области (предлагаемая при реформировании)

Решение общественных и внешних проблем эффективности и устойчивости здравоохранения предложено возложить на три Агентства. Задачей этих Агентств является обеспечение системы здравоохранения необходимыми базовыми ресурсами (информационные услуги, ценообразование, ресурсное и правовое обеспечение, организация лекарственной помощи). В соответствии с этими задачами Агентства могут взять на себя следующие неспецифические для органа государственной власти функции: исполнительно-распорядительные, оказания государственных услуг, управления государственным имуществом, обеспечивающие.

Исходя из этих принципов предложено создание следующих Агентств:

1. Агентство по экспертизе, стандартизации, управлению качеством и информационным технологиям здравоохранения Самарской области;
2. Агентство по ресурсному и правовому обеспечению здравоохранения Самарской области;
3. Самарское областное фармацевтическое агентство.

Прообразами Агентства по экспертизе, стандартизации, управлению качеством и информационным технологиям здравоохранения Самарской области и Агентства по ресурсному и правовому обеспечению здравоохранения Самарской области стали соответственно Самарский областной медицинский информационно-аналитический центр и Государственное предприятие «Перспектива», которые, начиная с 1991 года, последовательно брали на себя выполнение функций, обеспечивающих деятельность департамента и оказания государственных услуг (то есть де-факто уже являющиеся агентствами).

Суть разработанной структуры - скоординированная децентрализация управления, обеспечивающая последовательное выполнение принципа «управление на основе собственной компетенции». На высшем уровне управления принимаются стратегические решения, осуществляется законотворческая деятельность, планирование и распределение основных ресурсов. На следующих уровнях в базовых подразделениях принимаются самостоятельные оперативные управленческие решения. Разделение задач управления по уровням повышает качество и увеличивает эффективность управленческих решений.

Таким образом, предложенный нами вариант реформирования близок третьему сценарию административной реформы, изложенному в Докладе Высшей школы экономики «Насущные задачи реформы аппарата исполнительной власти», являющемуся, по мнению авторов, наиболее оптимальным. [2]

Список литературы:

- [1] Петр Щедровицкий. Единое управленческое пространство URL : <http://www.expert.ru>
- [2] Доклад Высшей школы экономики «Насущные задачи реформы аппарата исполнительной власти». Коммерсант № 25, 12 февраля 2004.
- [3] Отчет о научно-исследовательской и опытно-конструкторской работе Государственного университета-Высшей школы экономики Министерства экономического развития и торговли РФ «Разработка предложений и подготовка экспериментов по совершенствованию экономических механизмов деятельности государственных органов в рамках Концепции реформирования системы государственной службы Российской Федерации», шифр 9.20.6. УДК 33 № госрегистрации 01.200.2 07767 . Москва, 2002.

ТЕРРИТОРИАЛЬНЫЕ ОСНОВЫ СОВЕРШЕНСТВОВАНИЯ СИСТЕМЫ УПРАВЛЕНИЯ

Г.Г. Шалмина¹, В.В. Тарасевич²

¹Сибирская государственная геодезическая академия
Россия, 630108, Новосибирск, ул. Плеханова, 10
Тел. +7 (3832) 43-39-37, факс (3832) 44-30-60

²Новосибирский филиал РосНИИ Искусственного интеллекта
630090, Новосибирск, пр. Лаврентьева, 6, Россия
tvv@iis.nsk.su;
тел.: +7 (3832) 34-29-91, факс +7 (3832) 32-83-59

Ключевые слова: система, природная среда, население, социум, экономика, закономерность, недоопределенность, вычисления, критерий, управление

Abstract

The issues are considered concerning the management of complex territorial - public systems including the population with all kinds of its activity and the environment. It is asserted that it is necessary to take into account the laws of functioning of complex system: economic activities - the population - territory for correctly organized management. The account only the commodity-money parameters describing industrial and financial activity and not taking into consideration the conditions, interests and needs of the population, and also quality of an environment, results in one-sided and incorrect administrative decisions. The analysis of consequences of such erroneous decisions is given. The complex of efforts for development of correct effective decisions at the state and regional levels is suggested.

Введение

Социально-экономические системы относятся к наиболее сложным видам динамических систем. Их сложность обусловлена не только большим количеством и разнообразием элементов и связей, но и значительной степенью неопределенности, а также неформализуемостью многих зависимостей, что делает процесс принятия решений неформализованной, во многом интуитивно-эмпирической процедурой [1]. При выработке устойчивых управленческих решений, касающихся такой системы, для получения предполагаемого результата необходимо учитывать законы функционирования сложных систем [2]. Наличие так называемых странных аттракторов говорит о невозможности произвольного управления сложной системой, т.к. процесс неизбежно будет выходить на один из внутренне присущих системе сценариев развития событий [3]. Поэтому адекватное управление должно опираться на собственные свойства системы и учитывать их при постановке целей управления и выработки критериев.

Процесс принятия решения должен опираться на достаточно полную модель системы, иначе выработанное управленческое решение может оказаться неадекватным [4]. Однако, особенно в настоящее время, управленческие решения, как правило, затрагивают только административный аппарат управления и производственно-финансовую деятельность (экономику). Процесс принятия решений в большинстве случаев не затрагивает такие важные сферы как состояние социума во всех его аспектах, состояние природной среды, экологии среды жизнедеятельности и уровня хозяйственного развития территории. На наш взгляд, адекватная постановка вопроса невозможна без учета этих важнейших компонент эволюции не только человеческого общества, но и всего живого на Земле [5, 6]. Не удивительно, что такие управленческие решения, грешащие «экономизмом», приводят к декларированию целей, которые по сути недостижимы.

В настоящей работе предлагаются основы методологии управления сложными территориально-общественными системами (ОПС), более полно учитывающий все ее составляющие.

1 Цели и последствия управленческих решений

Во все времена независимо от политического строя основной задачей государства было и остаётся обеспечение материального благосостояния и социального благополучия граждан. Здоровье населения (духовное и физическое) - один из основных результатов функционирования государства. Если реформы в России направлены на создание условий необходимых для всестороннего развития человека как личности, то, по-видимому, следует считать анти-реформами: спад производства (более 50%); ухудшение здоровья населения (уменьшение средней продолжительности жизни “сильной” половины - до 54 лет); отрицательный баланс естественного прироста населения - (показатель рождаемость-смертность: -7 на 1000 человек); рост алкоголизма (в 2,5 - 3,0 раза); наркомании, самоубийств на фоне увеличения безработицы и обнищания населения (40-45% населения за чертой бедности) и т. д.

2 Социально-экономические циклы

Аналитический обзор процессов развития экономики в сочетании с динамикой форм государственной власти (XVII-XX столетия) позволил выявить чётко проявляющуюся цикличность процессов. Апогеи развития экономики, как правило, приурочены к периодам совпадения нескольких факторов, основными из которых являются: уровень подготовленности государственной системы и населения к следующему циклу (за 10-15 лет до начала перехода при средней продолжительности развития цикла 30-35 лет); усиление территориального аспекта управления; смена форм или введение новых парадигм управления; появление правителя-лидера, радеющего за судьбу государства и народа.

Решающее значение при усилении территориального аспекта системы имеют профессионализм руководства – знание, умение выявлять и учитывать при выборе стратегии управления, состояние объекта, особенности процессов его функционирования и развития.

3 Современная система управления территориальными частями России

Хотим мы того или нет, система управления Россией, как части объективного мира должна создаваться на основные составляющих этого мира, т.е. с учетом: состояния общества (развития социума, т.е. сообществ человека разумного); состояния природной среды (природно-климатических условий и имеющихся ресурсов); производной взаимодействия (общество-природа); состояние экологии среды жизнедеятельности.

За время «перестройки» многие элементы существовавшей системы управления утратили свое значение, а вновь созданные осложнили организационную структуру. На сохранившиеся фрагменты экономической отраслевой структуры (отрасли группы А и Б) наложены: схема, объединяющая отрасли материального и нематериального производства и схема, сформированная по принципу функциональной классификации отраслей (первичного и вторичного производства). Функционирует схема политико-административного «территориального» управления (три уровня - муниципальный, уровень субъектов федерации и федерации, включая структуру округов) и корпоративного управления.

Государство замкнуто на формировании аппарата администрирования, управляя социально-экономическими процессами опосредовано, через себе подобные структуры более низких рангов, представители которых, за редким исключением, профессионально не владеют ни технологиями производства, ни современными методами разработки стратегии управления территорией.

Главными недостатками современного управления являются: 1) громоздкость организационной структуры; 2) не четкое распределение функций (их наложение); 3) многоуровневое прохождение решений от центра к территории; 4) практическое отсутствие связи от территории к центру; 5) неразработанность механизма реформирования; 6) практически полное устранение центральных органов власти от решения социальных проблем населения.

В качестве основных средств управления выступают законодательно-правовые ограничения и только косвенно финансовое обеспечение через фискальную, бюджетную и внебюджетную политику финансирования.

Развитие социальных процессов на территории финансируется по сверхстаточному принципу. Финансовые ресурсы для обеспечения социально-экономических проблем территорий, по-прежнему, перераспределяются государством, несмотря на декларирование принципа местного самоуправления. Данный процесс осложняется практически 80% дотационностью муниципальных образований и 62% - субъектов федерации (налоговые отчисления консолидированного бюджета в бюджеты вышестоящих уровней возросли с 63% в 1995 до 92% в 2002 году), а ответственность за решение социальных проблем напротив перераспределена на уровень местного самоуправления. Кроме того, в социально-экономических программах развития территорий, акцентируется внимание на решении специфических отраслевых проблем, а проблемы социальные – жизнеобеспечения населения, как правило, не обсуждаются - «народ, как дитя выливается из купели».

При вновь грядущем перераспределении остатков государственной собственности (процессы приватизации), как показала практика проведения подобных акций, негативность социальных процессов (обнищание, деградация населения) возрастут многократно. Особенно остро встанет проблема создания соответствующего хотя бы средним мировым стандартам состояния среды жизнедеятельности населения. Сегодня не требует дополнительных доказательств положение о том, что здоровье и продолжительность жизни человека доопределяются состоянием этой среды.

4 Основы рациональной организации территориальных частей

4.1 Выбор совокупного критерия развития территории

При выборе вариантов развития территории необходимо опираться на тот или иной интегральный критерий. Наиболее распространенными до последнего времени были затратные критерии, в частности, приведенные затраты и дифференциальная рента [7].

В настоящей работе в качестве комплексного показателя состояния территории предлагается использовать совокупные затраты CZ_{iq} в виде [7]:

$$(1) \quad CZ_{iq} = \left(S_{iq}^C + S_{iq}^t + S_{iq}^0 + S_i^{\delta} + \sum V_{iq} + W_{iq}(l_q) \sum_n k_{iq} \cdot \varepsilon + d_{iq} \right) \cdot \eta_{iq} \rightarrow \min ,$$

где i – индекс вида сырья; q – индекс вида технологии; S^C – себестоимость без амортизационных отчислений; S^t – транспортные затраты; S^0 – плата за природопользование и санкции за загрязнение окружающей среды; S^{δ} – текущие затраты на обеспечение безопасности жизнедеятельности; $\sum V_{iq}$ – сумма всех видов налогов за исключением налогов на прибыль; W – корректировочный коэффициент, учитывающий влияние среды; l – координата размещения производства; $\sum k_{iq}$ – значение капитальных затрат производства (суммируются по всем звеньям производства); ε – коэффициент приведения капитальных затрат; d – затраты на продвижение товарного сырья на рынок; η – коэффициент неучтенных затрат.

Предлагаемый показатель развития территориальных процессов учитывает все основные характеристики составляющих ОПС комплексов (в обязательном порядке оцениваются специализирующие для данной территории виды деятельности) и рассчитывается в полном объеме или удельном исчислении (на единицу сырья, продукта). Он, по сути, отражает не менее 85-

90% предстоящих денежных затрат, при получении которых должны быть учтены количественные, полуколичественные и качественные характеристики составляющих территории, и отображает объем минимально необходимых затрат для реализации направления, которое можно условно быть принято за нижний уровень формирования цены реализации товарного сырья и продукта, получаемых на выходе конкретной рассматриваемой технологической цепочки [7].

4.2 Методика определения управленческого статуса территориальных частей

Критерием отнесения территориальной части к тому или иному управленческому уровню является показатель суммарного значения факторной нагрузки

$$(2) \quad \Phi H_m = \sum_{j=1}^{k\Phi} \alpha_j^m P_j^m / P_{n,j}^m,$$

где m – уровень структуры управления; $k\Phi$ – количество факторов или функций, вводимых в оценку; P_j^m – значение j -того фактора управления на m -том уровне управления; $P_{n,j}^m$ – нормативное значение показателя фактора или функции для m -того уровня управления; α_j^m – весовой коэффициент значимости j -того фактора или функции для m -того уровня управления.

Оценка завершается сравнением территориальных частей по удельному значению факторной нагрузки, приходящейся на одного жителя $\Phi H_m / N_S$, или приходящейся на единицу площади, занимаемой этим образованием, $\Phi H_m / S$, где S – размер площади территориальной части; N_S – численность населения, проживающего на этой территориальной части.

Окончательное решение о структурировании территории принимается на основе совмещения результатов факторной и функциональной оценок.

4.3 Распределение консолидированного бюджета на социальные нужды

Решение социальных проблем населения территориальных частей предлагается на основе распределения социальной доли государственного (в том числе территориального) консолидированного бюджета KB .

При этом по каждому j -тому фактору должны выполняться следующие ограничения

$$(3) \quad KB \cdot \sum_m k_{mj} N_m \geq a_j N,$$

где k_{mj} – доля расхода консолидированного бюджета на m -том уровне управления по j -тому фактору; N_m – число жителей в границах данной ТЧ, относящейся к m -тому уровню управления, N – общее число жителей; a_j – норматив бюджетной обеспеченности одного жителя по j -тому фактору.

Обеспечение предполагаемого социального уровня жизнедеятельности населения определяется следующим образом:

$$(4) \quad KB \cdot k_{sm} \geq a_{\min}^{соц},$$

где k_{sm} – социально-гарантированный прожиточный минимум; $a_{\min}^{соц}$ – часть расходов населения для обеспечения необходимого жизненного уровня.

В качестве целевой функции будет использоваться функция эффективности работы аппарата управления $f_{эу} = AA/KB$:

$$(5) \quad f_{эу} \rightarrow \min,$$

где AA – затраты на содержание управленческого аппарата, при соблюдении условия:

$$(6) \quad \partial f_{\text{эв}} / \partial t \leq 0.$$

Последнее условие означает, что величина AA должна находиться в определенной зависимости от уровня финансовой обеспеченности населения.

В рамках указанной постановки задачи можно:

- при заданных долях k_{mj} определить минимальную величину KB , удовлетворяющую соотношениям (3) – (6);
- при заданной величине KB найти доли k_{mj} , удовлетворяющие (3) – (6).

Может оказаться, что при заданном уровне KB решений не существует. В этом случае максимальное нарушение неравенств (3) – (4) будет показывать наличие узких мест в социальной политике государства – необходимость увеличения KB до тех пор, пока результаты решения задачи не приблизятся к реальному уровню решения социальных проблем.

4.4 Методика размещения производства и рациональной организации базисной части системы

Как уже отмечалось выше, принятие решений, касающихся развития территориальной общественно-природной системы и рациональных организации и размещения всех видов деятельности населения, оперирует параметрами и зависимостями, многие из которых обладают значительной степенью неопределенности, расплывчатости, и трудно поддаются количественным оценкам и формализации в виде аналитических зависимостей. Оцифровка таких параметров носит во многом волонтаристский характер и не способствует получению адекватных оценок. Более перспективным представляется здесь использование нетрадиционной математики, например, теории расплывчатых множеств [8] или аппарата недоопределенных вычислений [9, 10], позволяющих работать с неточно и/или неполно заданной информацией и обладающих развитыми средствами для отображения качественных зависимостей.

В работах [1, 7] рассмотрена идея методики последовательного размещения технологических цепочек и звеньев производства на территории. Наряду с этим, может быть предложена методика рационального размещения производства одновременно как основного, так и побочного продукта с учетом поступления отходов в окружающую среду [1, 11], и других видов деятельности, на основе аппарата недоопределенных вычислительных моделей. Семантическая сеть строится исходя из существующих и планируемых вариантов размещения производительных сил, при этом параметры размещаемого производства (географическая привязка, величина и направление транспортных потоков и т.п.) являются недоопределенными величинами, а география функционирующих объектов известна. Семантическая сеть содержит как внутренние вычислительные модели, так и глобальные, характеризующие различные (технологические, экологические, финансовые, хозяйственные и т.п.) ограничения и связи между объектами. В процессе активации семантической сети происходит отбор вариантов по максимальному значению общей валовой прибыли и чистого дисконтированного дохода. Процесс вычислений завершается сужением пространства выбора до одного или нескольких вариантов с минимальным значением критерия (1). Одновременно уточняются географические координаты производства и его основные параметры.

По завершении выбора варианта развития территории определяется оптимальная структура управления путем выбора вариантов, удовлетворяющих ограничениям (2) – (6).

Заключение

Современные социально-экономические процессы в России не только не соответствуют ее геополитическим интересам, а скорее противоречат им. Сохранение режима суженного воспроизводства в России стало свершившимся фактом. Проблема деградации населения требует коренного пересмотра подхода к ее решению со стороны государства. По мнению авторов статьи, необходимо перейти от декларативного подхода к решению социально-экономических

проблем территории к объективно обусловленным решениям на государственном и региональном уровнях: во-первых, провести комплексное социально-экономическое зонирование территориальных частей России (паспортизацию хозяйства, природной среды, экологии, здоровья населения, уровня жизни и т.д.); во-вторых, разработать единые территориальные основы выбора стратегии развития экономики; в-третьих, осуществить реформирование управления с учетом состояния территориальных частей страны, в качестве обязательной основы совершенствования экономики и развития страны как единой общественно-природной системы [12].

Использование аппарата недоопределенных вычислений и других технологий искусственного интеллекта снимает многие ограничения и трудности, характерные для применения классических методов расчет экономических оптимизационных задач, и предоставляет новые возможности для описания и анализа ОПС.

Список литературы

- [1] Шалмина Г.Г., Тарасевич В.В., Загарин А.В. Основы стратегии развития регионов России. – Новосибирск: НГАЭиУ, 1999. – 420 с.
- [2] Князева Е.Н., Курдюмов С.П. Законы эволюции и самоорганизации сложных систем. – М.: Наука, 1994. – 238 с.
- [3] Белавин В.А., Курдюмов С.П. Уникальность моделей сложных систем // в сб-ке "Математическое и компьютерное моделирование социально-экономических процессов". – Нарофоминск, декабрь 2000г., стр.38-48.
- [4] Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. – М.: Высш. шк., 1989. – 367 с.
- [5] Курдюмов С.П. Законы коэволюции социальных систем человечества и природы // в сб-ке "Всероссийский форум «Миллион друзей». Живая природа и общество". – Нижний Новгород, 2001г., с.10-14.
- [6] Шугрин С.М. Космическая организованность биосферы и ноосферы. – Новосибирск: Сиб. предприятие РАН «Наука», 1999. – 496 с.
- [7] Шалмина Г.Г. Основы экологического менеджмента. – Новосибирск: Сибирская академия государственной службы, 2002. – 260 с.
- [8] Нечеткие множества в моделях управления и искусственного интеллекта. /под ред. Д.А.Поспелова. – М.: Наука, 1986.
- [9] Напреенко В.Г., Нариньяни А.С. и др. Моделирование национальной экономики с использованием недоопределенных моделей. // Проблемы управления и моделирования в сложных системах: Труды II Международной конференции. – Самара, Самарский научный центр РАН, 2000 г., с. 35 – 42.
- [10] Напреенко В.Г., Нариньяни А.С., Юртаев А.В. Недоопределенные модели – нетрадиционный подход к математическим исследованиям экономики. – Информационные технологии, 1999, № 4, с. 36-41.
- [11] Шалмина Г.Г., Тарасевич В.В. Сетевые недоопределенные модели пространственной организации территориальной общественно-природной системы //Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции. – Самара, Самарский научный центр РАН, 2002 г., с. 35 – 42.
- [12] Территориальные основы государственного управления. В 2-х ч. Часть II. Системный подход к разработке территориальных основ государственного управления. /под ред. Г.Г.Шалминой. – Новосибирск, Изд-во НГОНБ, 2003. – 256 с.

РАЗРАБОТКА ПЕРВОЙ ОЧЕРЕДИ СИСТЕМЫ УПРАВЛЕНИЯ РЕГИОНОМ С ПРИМЕНЕНИЕМ МУЛЬТИАГЕНТНЫХ ТЕХНОЛОГИЙ

В.А. Виттих¹, Г.Д. Светкина², П.О. Скобелев³, Д.В. Волхонцев³, Е.А. Гриценко²,
А.Н. Никитин³, О.Л. Сурнин³, М.А. Шамашов³

¹Институт проблем управления сложными системами РАН
443020, Россия, Самара, ул. Садовая, 61
vittikh@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²Министерство труда и социального развития Самарской области
443054, Россия, Самара, ул. Революционная, 44
depart@socio.santel.ru

тел./факс: +7 (8462) 34-33-00

³ООО «НПК «Генезис знаний»
443011, Самара, ул. Ново-Садовая, 221, оф. 206, Россия
agro@kg.ru

тел./факс: +7 (8462) 34-87-57

Ключевые слова: социальный паспорт, персональный учет населения, базы данных, базы знаний, управление регионом, мультиагентные системы

Abstract

The realized new approaches to development of information systems for management of region are described. The key moment is maintenance of address interaction of each person with enforcement authorities on the basis of the concept of the social passport which can be used for construction of systems of the personal account of the population.

1 Создание системы управления регионом

Управление регионом – сложная задача, основная цель которой заключается в создании условий для расширения возможностей каждого человека реализовывать свои устремления, вести здоровую, полноценную творческую жизнь, расширять возможности интеллектуального, политического, социального и экономического выбора.

В настоящее время государство воздействует на человека, заставляя его приспосабливаться к бюрократической системе. Суть же системы управления состоит в замене логики воздействия принципом взаимодействия. В общем случае, система управления охватывает три ветки взаимодействия – это взаимодействие органов власти с другими органами исполнительной власти, взаимодействие власти с населением и взаимодействие власти с бизнесом.

Создание, внедрение и эксплуатация первой очереди системы управления регионом с применением мультиагентных технологий невозможны без реформирования и адекватных изменений государственной службы, без информационной и компьютерной поддержки этих реформ и, в первую очередь, в социальной сфере.

Вся деятельность социального блока должна стать открытой и направленной на адресное взаимодействие и помощь гражданам региона. При этом централизованное управление должно заменяться на децентрализованное, иерархическая вертикальная структура уступит место сетевой и горизонтальной, инструкции «сверху вниз» – на переговоры равных сторон, управление по затратам – на управление по проектам, работа с кадрами – на развитие потенциала человека, оплата структур управления – на оплату функций и результатов и т.д.

Для того, чтобы сфокусировать на проблемах каждого конкретного человека деятельность региональных органов власти, представительств органов федеральной власти и муниципальных образований, необходимо создать соответствующие интеграционные механизмы, в основу которых предложено положить концепцию социального паспорта жителя Самарской области [1].

Повышение качества принимаемых управленческих решений напрямую связано с обеспечением адресного взаимодействия органов исполнительной власти с каждым жителем региона (например, адресная социальная поддержка и медицинская помощь, учитывающие индивидуальные особенности и социальное положение человека, возможность самостоятельного выбора траектории образования и т.д.). Обоснованием всех принимаемых управленческих решений в департаментах социального блока должны выступать реальные запросы граждан. Причем, важной особенностью системы управления регионом должна стать её направленность на решение социально-значимых функций, а не на автоматизацию деятельности отдельных ведомств и организаций. Таким образом, «первый этаж» разрабатываемой «двухъярусной» системы будет реализовывать адресное взаимодействие органов власти с населением, а «второй этаж» - непосредственно осуществлять управление департаментами социального блока, включая мониторинг потребностей и возможностей населения, выработку вариантов решения и контроль эффективности их исполнения.

2 Принципы построения системы

В реализованной системе процессы управления реализуются путем сочетания иерархических («вертикальных») и сетевых («горизонтальных») методов принятия решений. Это означает, что проблемы, возникающие в регионе, решаются на основе многосторонних взаимодействий элементов системы, к которым относятся не только органы государственной власти и местного самоуправления, но и хозяйствующие субъекты, и жители Самарской области. Данный подход соответствует принципам проводимой в настоящее время административной реформы.

Адресное взаимодействие человека с государственными структурами обеспечивается менеджером по работе с населением в соответствии с требованиями реформы государственной службы. Менеджер по работе с населением работает с гражданами, прикрепленными к определенной территории, на которую в пропорциональном отношении отпускаются ресурсы (в том числе и финансовые) для предоставления сервисов населению, а менеджер является распорядителем этих ресурсов в соответствии с потребностями населения. Взаимодействие менеджера с гражданами определяется на основе регламентов взаимодействия, которые разрабатываются управленческим звеном более высокого уровня – менеджером по организации процессов управления. Менеджер по организации процессов управления контролирует деятельность менеджеров, работающих с населением.

Одним из принципов построения системы является использование баз знаний, представленных в виде семантической сети понятий и отношений, характерных для социальной сферы. Через семантическую сеть экземпляров этих понятий и отношений представляются конкретные ситуации каждого гражданина, например, персональные данные человека, извлекаемые из внешних баз данных, к которым применяются хранящиеся в базе знаний правила рассуждений и принятия решений. Важной особенностью является интеграция в едином месте федерально-регионального и муниципального законодательства.

3 Архитектура системы

Первая очередь системы управления регионом с применением мультиагентных технологий включает три основные функциональные подсистемы, способные взаимодействовать как между собой, так и напрямую с жителями Самарской области через сеть Интернет или Интернет-киоски:

- систему адресного взаимодействия населения и органов исполнительной власти в социальной сфере;
- систему поддержки принятия решений для управления департаментом социального блока;
- открытый Интернет-портал для интеграции ресурсов и поддержки межведомственного взаимодействия.

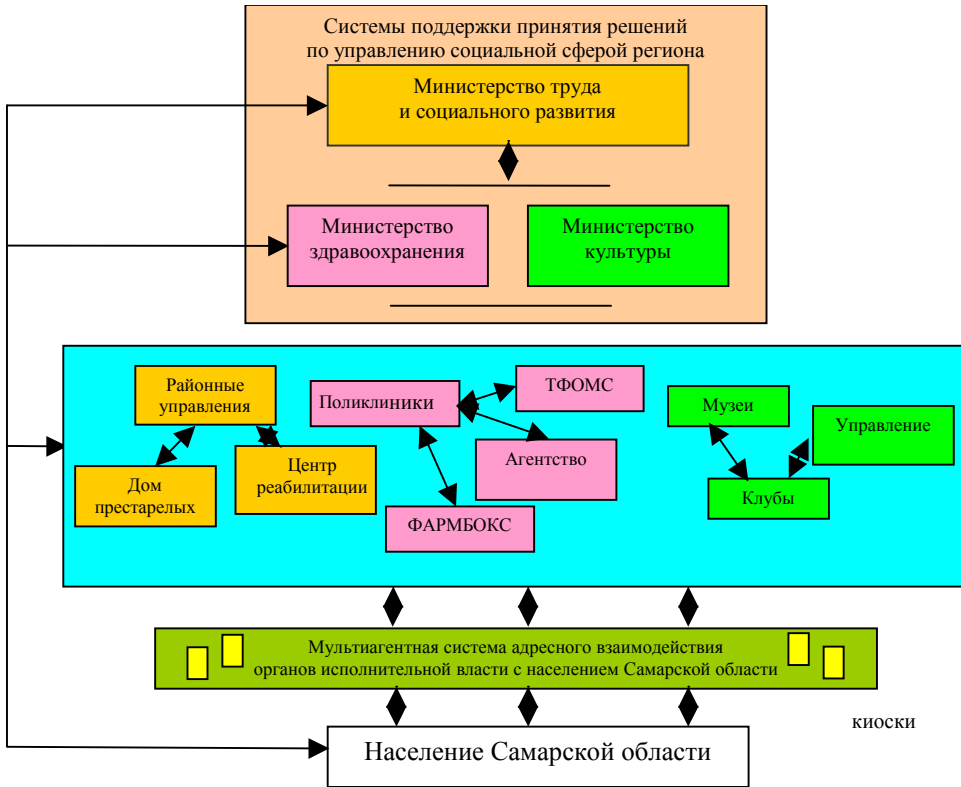


Рисунок 1 – Архитектура системы

Такое разделение на подсистемы определяется во многом автономными и самодостаточными функциями этих подсистем, ориентированными на различные целевые группы пользователей, и тремя базовыми направлениями развития предлагаемой системы:

- государство для населения (G2C), представленного мультиагентной системой адресного взаимодействия населения и органов исполнительной власти,
- государство для государства (G2G), представленного системой поддержки процессов принятия решений для управления департаментом социального блока,
- государство для предприятий и бизнеса (G2B), представленного открытым Интернет-порталом для интеграции ресурсов и поддержки межведомственного взаимодействия.

3.1 Подсистема адресного взаимодействия населения и органов исполнительной власти

Система адресного взаимодействия населения и органов исполнительной власти на основе социального паспорта формирует первый (нижний) уровень системы управления регионом. Эта система позволяет адресно, т.е. на основе персонифицированной информации о человеке, найти федеральные, региональные и муниципальные законы социальной сферы (социальная защита, здравоохранение, культура), применимые к пенсионерам и ветеранам труда, инвалидам и другим категориям граждан, и активизировать междепартаментные регламенты взаимодей-

ствия человека с органами исполнительной власти для решения их насущных проблем. Кроме того, система позволяет человеку или менеджерам по работе с населением промоделировать различные ситуации, например, показать человеку, какие льготы будут для него действовать при изменении его социального статуса (например, если человек становится ветераном труда).

3.2 Подсистема управления департаментом социального блока

Мультиагентная система для управления департаментом социального блока предназначена для поддержки процессов принятия решений, связанных с оказанием услуг населению, в департаментах социального блока.

Подсистемы каждого департамента дают возможность наблюдать ситуации, связанные с конкретными людьми в подведомственных учреждениях, реализуемые на нижнем уровне (а также взаимодействием организаций), идентифицировать возникающие там проблемы, которые не могут быть разрешены на нижнем уровне, вырабатывать и согласовывать варианты междепартаментных решений и обеспечивать их выполнение (как, например, перевод человека на инвалидность или его реабилитация, недостаток финансовых средств, положенных по закону определенным категориям граждан и т.д.). Кроме того, на этом уровне обеспечивается поддержка процессов принятия решений по управлению ресурсами, предназначенными для предоставления услуг населению, например, когда обнаруживается, что не посещаемый клуб, поликлинику с малым количеством пациентов или дом престарелых в одном районе целесообразнее реорганизовать или открыть в другом районе, где, на основании анализа данных, имеющих в системе, будет гораздо более актуальна потребность в них.

3.3 Открытый Интернет-портал интеграции ресурсов и поддержки межведомственного взаимодействия

Интернет-портал для интеграции ресурсов и поддержки межведомственного взаимодействия позволяет менеджерам организаций социальной сферы создавать собственные «представительства» в системе и определять функции своих организаций и регламенты их взаимодействия, как с другими организациями, так и с гражданами, через описание потребностей и возможностей.

На этой основе могут быть построены отраслевые подсистемы управления, включающие в качестве организаций для департамента социальной защиты - районные управления, социальные и реабилитационные центры, для департамента здравоохранения - поликлиники, территориальные отделения фонда медицинского страхования (ТФОМС) и уполномоченные региональные компании по снабжению льготными лекарствами (ФАРМБОКС) и т.д. (рисунок 1). При идентификации человека через подсистему адресного взаимодействия, с его разрешения, указанные организации получают возможность делать предложения человеку и взаимодействовать для удовлетворения его потребностей. Например, временный переселенец будет приглашен в конкретную поликлинику или медицинский центр, больному инсулином после постановки диагноза сразу придет предложение центр «Диабет», на новую выставку икон древней Руси в художественном музее будут приглашены классы школ, изучающие по программе данную тему, поликлиника сможет напрямую взаимодействовать с ФОМС по вопросам инвестиций для приобретения оборудования и т.д.

4 Социальный паспорт в системе персонального учета населения

Под «социальным паспортом» чаще всего понимается система важнейших показателей организации или целого региона, документально отображающих перспективы социального развития, социальный потенциал. Эти данные представляют только общие тенденции общественно-экономического развития на определенный момент, за ними не видно конкретного человека с его проблемами. Кроме того, эта статистика не может служить основой для понимания

каждым жителем его прав и возможностей, доступности тех или иных социальных льгот и услуг, предоставления информации об уже оказанной помощи.

В контексте настоящего проекта социальный паспорт – это совокупность социально-значимых сведений о человеке, определяющих его общественный, имущественный и правовой статусы, полученных в соответствии с действующими нормативными актами, хранящихся в электронной форме и доступных как самому человеку, так и всем организациям, оказывающим ему тот или иной вид услуг [2].

Такой подход к пониманию понятия социального паспорта:

- позволяет реализовать принцип «единого окна», чем облегчает процедуру оформления и предоставления различных видов социальной помощи;
- служит основой для назначения единого социального пособия, имеющего адресную направленность;
- обеспечивает предоставление информационных услуг населению;
- является основой для создания системы принятия решений в сфере социальной политики;
- позволяет осуществлять учет и контроль расходуемых средств вплоть до определения льгот и услуг, оказанных конкретному получателю.

Социальный паспорт, создаваемый для каждого жителя Самарской области, содержит персонифицированную информацию, характеризующую его социальный статус, семейное и материальное положение, социальные льготы и т.д. Для хранения этой информации должны быть использованы территориально распределенные информационные базы данных, которые будут использоваться органами исполнительной власти для выработки индивидуальных программ взаимодействия с каждым гражданином.



Рисунок 2 – Карта социального паспорта

Доступ к информации социального паспорта предоставляется человеку посредством микропроцессорной пластиковой карты-ключа (рисунок 2), с помощью которой человек может активировать любой Интернет-киоск, расположенный для свободного доступа в общественно-значимых местах на территории Самарской области (учреждениях исполнительной власти, поликлиниках, социальных учреждениях и т.д.) и получить санкционированный доступ к персональной информации. Память пластиковой карты содержит только краткие сведения о человеке и ключи доступа к его персональной информации в существующих ин-

формационных ресурсах. На время сеанса обращения человека к системе (пока карточка физически находится в считывателе Интернет-киоска) образуется временная структура данных, которая по завершении работы удаляется для обеспечения безопасности.

Таким образом, организация адресного взаимодействия населения и органов исполнительной власти осуществляется на основе концепции социального паспорта жителя Самарской области как совокупности социально-значимых сведений о человеке, определяющих его общественный, имущественный и правовой статусы, хранящихся в электронной форме в существующих информационных ресурсах органов исполнительной власти. Социальный паспорт аккумулирует в себе всю существующую персональную информацию о человеке только на период обращения человека к системе, при этом центральная база данных системы содержит только ключевые ссылки на персональную информацию, которая хранится во внешних информационных ресурсах. Подобная реализация проекта выполнена в соответствии с принципами создания системы персонального учета населения и позволяет в рамках первой очереди системы управления регионом реализовать пилотную версию системы персонального учета населения.

5 Принципы построения системы персонального учета населения на базе социального паспорта

При создании (на основе концепции социального паспорта) системы персонального учета населения реализуются следующие основные принципы:

- 1) Обеспечение доступа к персональной информации в соответствии действующим законодательством.

Доступ к персональной информации осуществляется исключительно в соответствии с правовыми актами уровня федерального законодательства.

- 2) Полномочность ведомств в части обеспечения функционирования собственных информационных систем.

Органы государственной власти и местного самоуправления самостоятельно управляют процессами сбора, хранения, внутриведомственного использования и уничтожения информации, включая вопросы обеспечения защиты информации от несанкционированного доступа. Система персонального учета населения является лишь средством интеграции ведомственных информационных систем, позволяющим достичь нового качества информации за счет ее совместного использования.

- 3) Доступность для граждан персональной информации о себе.

В рамках системы обеспечивается свободный доступ граждан к собственной персональной информации (кроме случаев, когда доступ к такой информации ограничен законодательством). При функционировании системы обеспечивается протоколирование фактов доступа к персональной информации, и исполненного административного регламента.

- 4) Наличие общего ключевого поля во всех информационных системах номера карты социального паспорта.

В рамках системы обеспечивается однозначная сопоставимость данных в различных информационных системах за счет введения в них единого уникального ключевого параметра (поля) – номера карты социального паспорта.

- 5) Расширяемость системы.

В системе учитывается принцип расширяемости за счет присоединения к системе внешних информационных источников, содержащих персональную информацию.

Список литературы

- [1] В.А. Виттих, Д.В. Волхонцев, Е.А. Гриценко, Г.Д. Светкина, П.О. Скобелев. Разработка мульти-агентной системы для моделирования процессов взаимодействия населения и органов исполнительной власти в социальной сфере // Труды IV Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-24 июня 2002. – Самара: СНЦ РАН, 2002, с. 414-423.
- [2] В.А. Виттих, Д.В. Волхонцев, Е.А. Гриценко, А.Н. Никитин, Г.Д. Светкина, П.О. Скобелев. Социальный паспорт как распределенная база знаний и данных. // Труды V Международной конференции по проблемам управления и моделирования в сложных системах. Самара, 17-22 июня 2003. – Самара: СНЦ РАН, 2003, с. 386-390.

СЕТЕВАЯ МУЛЬТИАГЕНТНАЯ МОДЕЛЬ СИСТЕМЫ УПРАВЛЕНИЯ ЗДРАВООХРАНЕНИЕМ РЕГИОНА И СИСТЕМА КОНТРОЛЯ ЭФФЕКТИВНОСТИ И КАЧЕСТВА РАБОТЫ ВРАЧЕЙ ПОЛИКЛИНИКИ

С.В. Батищев¹, В.А. Виттих², Д.В. Волхонцев³, Г.И. Гусарова⁴, Г.П. Котельников⁵,
С.И. Кузнецов⁴, В.В. Павлов⁴, П.О. Скобелев¹, О.Л. Сурнин³, Л.С. Федосеева⁶, М.А. Шамашов³,
И.С. Шапиро³, Е.В. Чернов³

¹НПК «Маджента Девелопмент»

²Институт проблем управления сложными системами РАН

³НПК «Генезис знаний»

443011, Самара, ул. Ново-Садовая, д. 221, оф. 206

shamashov@kg.ru

тел/факс: 8 (8462) 34-87-57, 17-30-69

⁴Министерство здравоохранения Самарской области

⁵Самарский государственный медицинский университет

⁶Муниципальное медицинское учреждение городская поликлиника № 15

Ключевые слова: сетевая модель, холистический подход, мультиагентные технологии, онтология, врач общей практики, инвестиции, фармэкономика, интеграция, экономическая модель, медико-экономические показатели

Abstract

Consideration is given to a network multi-agent model aimed to support decision-making procedure in regional health care management. The model consists of prototypes of automated workplaces of a general practitioner and the chief medical officer, specialists of a hospital, ambulance center, Medical University pharmacological company, managers of the Compulsory Medical Insurance Fund and the Ministry of Health Care of the Samara District area. Characteristic examples are cited illustrating the interaction of software smart agents of the representatives of different health-care institutions while solving typical problems on evaluating the need of investments to the outpatients' clinic equipment, creation of a new medical product, solution of pharmacological economics issues, constant monitoring of a patient's state of health and cross-performance or integration of general practitioners, ambulance and hospital professionals. A sub-system is also described that is able to control a set of indices in order to make an estimation of the effectiveness and quality of general practitioners' performance at an out-patients' clinic.

1 Состав и назначение сетевой мультиагентной модели системы управления здравоохранением Самарской области

Сетевая мультиагентная модель системы управления здравоохранением региона реализована в ходе выполнения проекта «Разработка новой модели управления здравоохранением Самарской области». Этот проект стал одним из победителей второго раунда конкурса программы «Партнер» («Партнерство: взаимодействие, укрепление, развитие»), который проводился в Приволжском федеральном округе, под эгидой Международного Совета по Научным Исследованиям и Обменам (АЙРЕКС) и получил грант, финансируемый Агентством США по международному развитию (АМР США).

В задачи проекта входила не только разработка базовых направлений реформирования здравоохранения [1], но и построение компьютерной модели для поддержки процессов принятия решений при взаимодействии врача и пациента, учреждений здравоохранения Самарской области с населением на основе новой концепции управления здравоохранением. Эта сетевая

модель, отражающая новую холистическую организацию системы управления здравоохранением, является прообразом открытой мультиагентной системы (ОМАС) поддержки процессов принятия решений (ППР) “Регион” [2]. Без данной мультиагентной системы, способной обеспечивать взаимодействие организаций здравоохранения, самостоятельно находить варианты и принимать индивидуальные решения для сотен, тысяч и миллионов людей, обладающими множеством персональных особенностей, генерируя предложения по индивидуальным схемам диагностики, лечения и профилактики каждого человека, внедрение предлагаемой модели управления здравоохранением было бы просто невозможным.

Мультиагентные системы в сфере здравоохранения призваны обеспечить интеллектуальную поддержку процессов принятия решений в следующих областях:

- организация лечебно-профилактической работы,
- планирования ежедневной работы врачей и менеджеров,
- управления персоналом,
- управления финансами.

В предлагаемой мультиагентной модели, а затем и системе управления здравоохранением региона программный агент вместе со своей онтологией и персональными настройками может быть поставлен в соответствие:

- каждому пациенту и врачу,
- руководителю поликлиники и другим менеджерам лечебно-профилактических учреждений,
- любым ресурсам, таким как койка стационара, новое лекарство, метод лечения.

Агент каждого больного, действуя от имени и по поручению своего реального представителя, будет находить лучшие методы лечения, лекарственные средства, необходимые льготы, лучшего врача, специализирующегося на данном заболевании и т.д.

Агенты ресурсов, агенты возможностей призваны искать заинтересованных в них агентов потребностей и контролировать эффективность собственного использования.

Реализованная модель закладывает основу автоматизированных рабочих мест (АРМ) интегрированной мультиагентной системы для врачей общей практики (ВОП) и главного врача поликлиники, АРМ специалистов стационара, скорой помощи, медицинского университета, АРМ менеджеров Министерства здравоохранения региона, территориального фонда обязательного медицинского страхования (ТФОМС), уполномоченного фармакологического предприятия “ФАРМБОКС” и, наконец, любого гражданина – пациента медицинских учреждений, зарегистрированных в открытом мультиагентном медицинском портале Самарской области (см. рисунок 1).

Основа такого портала также была реализована в рамках рассматриваемого проекта. Основное его назначение состоит в интеграции и широком распространении профессиональных медицинских, управленческих и хозяйственных знаний, необходимых для работы сетевой модели и ОМАС ППР “Регион”. Эти динамические, пополняемые базы знаний (онтологии), представленные в портале в форме семантических сетей, включают в свой состав:

- Описание клинических протоколов, международных стандартов диагностики, лечения и профилактики, основанных на доказательной медицине.
- Персональные описания врачей и их лечебно – профилактического опыта.
- Истории болезни пациентов и их предпочтения.
- Описание современных лекарственных средств.
- Регламенты взаимодействия врачей и пациентов, медицинских учреждений, специалистов поликлиники между собой (управленческие и хозяйственные знания).

Медицинские базы знаний включают в свой состав не только схемы лечения, но и оценки стоимости, что позволяет персонализировано планировать бюджет и контролировать его исполнение.

Онтологии диагностики, лечения и профилактики в рамках системы могут выступать в качестве:

- навигатора врача,
- средства контроля качества лечения,
- интеллектуального консультанта врача и пациента,
- средства обучения.

Онтологии мультиагентной системы отделены от кода программы, ориентированы на конечного пользователя и не требуют специальных знаний в области информационной техники. Предусматриваемый в комплекте с системой конструктор онтологий будет ориентирован на специалистов в области здравоохранения и позволит им самостоятельно дополнять методики лечения различных заболеваний.

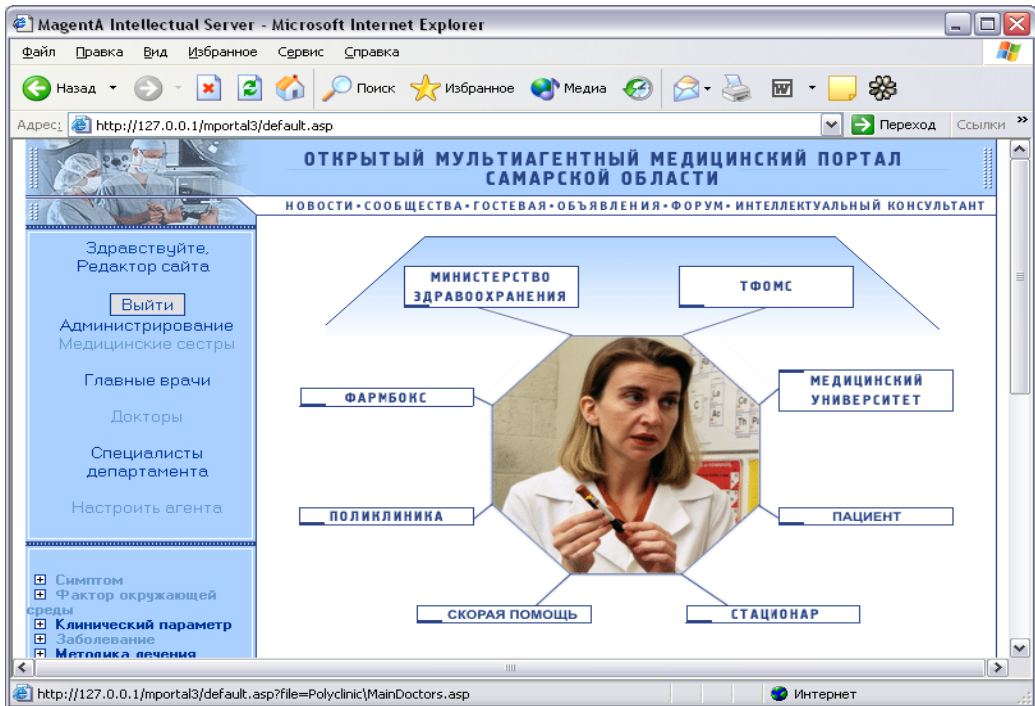


Рисунок 1 – Страница медицинского портала для выбора автоматизированных рабочих мест

2 Примеры моделируемых ситуаций взаимодействия организаций и специалистов здравоохранения

Реализованная сетевая мультиагентная модель позволила отработать новые принципы управления здравоохранением региона на типовых примерах ситуаций взаимодействия между организациями и специалистами здравоохранения. Рассмотрим несколько примеров моделируемых ситуаций, демонстрирующих эффективность использование мультиагентных технологий.

2.1 Принятия решений по инвестициям средств на приобретение оборудования

Рассмотрим ситуацию, которая начинается приемом пациента врачом общей практики. На своем рабочем месте ВОП отмечает начало нового случая и ставит диагноз – одна из разновидностей астмы.

Агент врача выбирает из онтологии, представляющей международный стандарт лечения бронхиальной астмы, схемы лечения данного заболевания, сортируя их по эффективности и возрастианию максимальной цены лечения, отмечая недоступные пока схемы. Список методик лечения выводится на страницу АРМ ВОП (рисунок 2). Врач не может выбрать наиболее эффективную рекомендуемую методику первого выбора из-за отсутствия в поликлинике необходимого оборудования и выбирает первую доступную методику (вторую в списке).

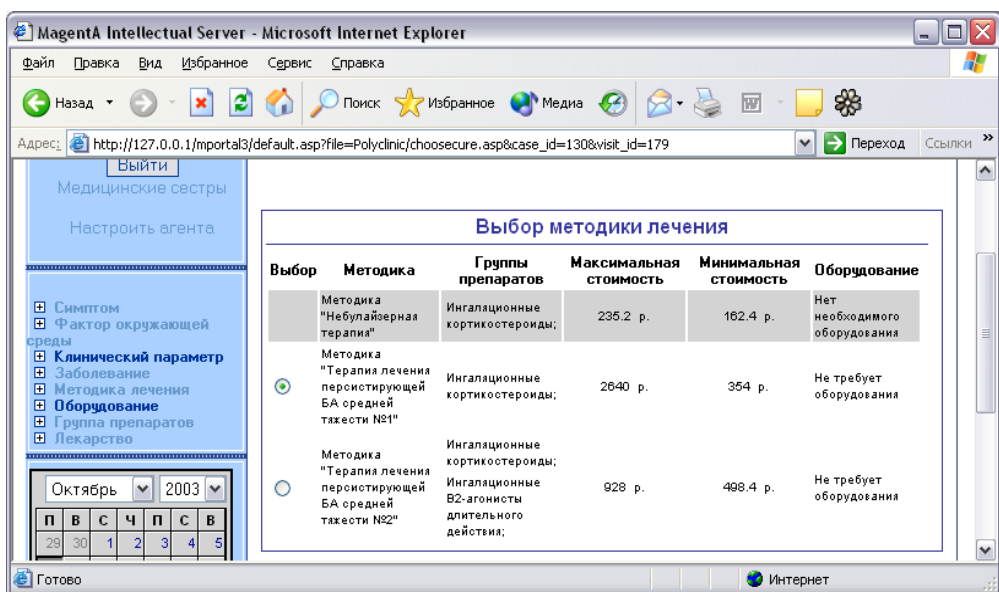


Рисунок 2 – Выбор методики лечения

Затем врач выбирает лекарственные препараты, соответствующие методике лечения, и с помощью своего агента, который извлекает из онтологии расписание работы врача и план приема пациентов, назначает пациенту дату очередного визита.

Агент главврача обнаруживает неэффективный выбор и определяет его причину (нет оборудования). Он рассчитывает минимальные потери поликлиники, возникающие при назначении методики, и сообщает об этом главврачу, с предложением подготовить письмо для ТФОМС и Департамента (Министерства). Сообщения об этом появляются на рабочем месте главврача (рисунок 3).

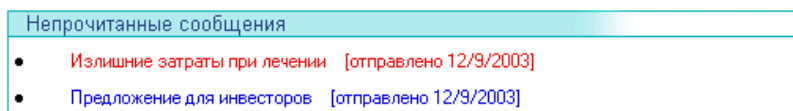


Рисунок 3 – Сообщения агента главного врача

Если главврач согласен на подготовку письма, то агент главврача готовит письмо в ТФОМС с просьбой выделения средств на данное оборудование, рассчитывая срок окупаемости в зависимости от стоимости оборудования и применимости методики (количества больных, к которым она может быть применена). При получении согласия главврача на отправку письма, агент главврача начинает переговоры с агентом менеджера ТФОМС о необходимых инвестициях. Схема переговоров агентов врача, главного врача, агентов менеджеров ТФОМС и Департамента здравоохранения представлена на рисунке 4.

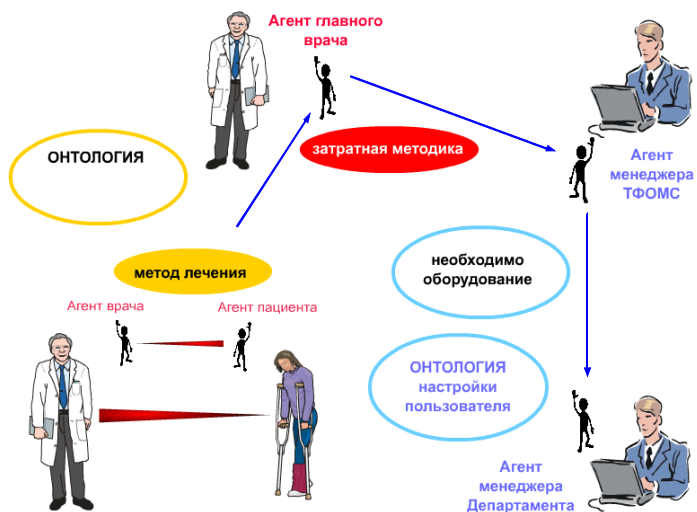


Рисунок 4 – Переговоры агентов о необходимости приобретения оборудования

Агент менеджера ТФОМС в зависимости от настроек предпочтений и ограничений (отказать в этом месяце и запланировать на будущее; выделить средства; выделить средства, если объем инвестиций и срок окупаемости удовлетворяет заданным условиям) оценивает, заслуживает ли внимание это предложение. Если агент менеджера ТФОМС согласен, то аналогичные переговоры повторяются между агентом менеджеров ТФОМС и Департамента. Если все согласны выделить средства, то агенты обращаются к своим “хозяевам” – менеджерам ТФОМС и Департамента (рисунок 5), и в случае их согласия выделить средства положительный ответ идет главврачу.

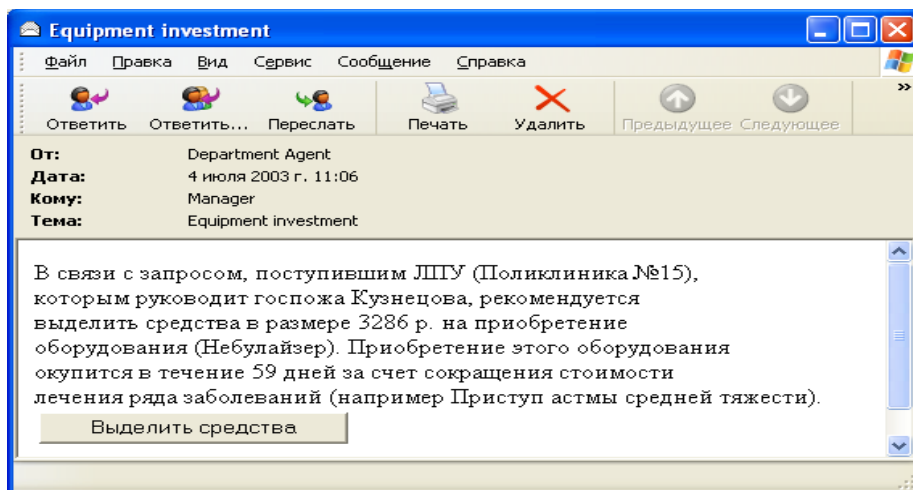


Рисунок 5 – Предложения об инвестициях

2.2 Расчет средств на лечение хронического больного

Предположим, что врач закрывает случай для пациента с окончательно установленным диагнозом бронхиальная астма, выбрав методику лечения из клинического протокола и назначив ему необходимые льготные лекарства. Агент врача на основании данных из клинического

протокола рассчитывает стоимость ежемесячного амбулаторного лечения выявленного хронического больного, в том числе и ежемесячную стоимость, необходимых ему льготных лекарств.

Агент врача доводит сведения о выявленном хроническом больном и стоимости его лечения до агента главврача, а тот, в свою очередь, до агентов менеджеров ТФОМС и Департамента (Министерства) здравоохранения. Агенты менеджеров, в зависимости от настроек, сообщают эту информацию заинтересованным специалистам (рис. 6.17) и заносят информацию о хроническом больном и стоимости его лечения в базы данных ТФОМС для формирования обоснованного плана выделения средств на лечение и льготные лекарства не на основе среднестатистических показателей, а персонифицированного учета.

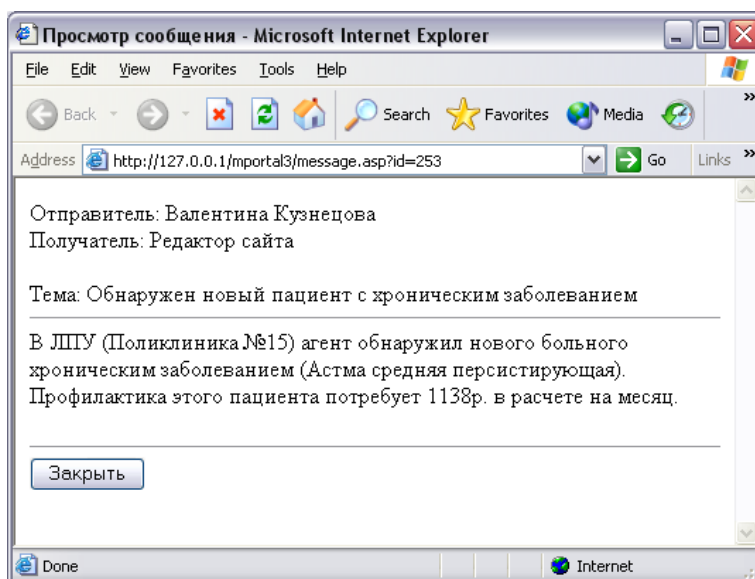


Рисунок 6 – Сообщение менеджеру ТФОМС о выявлении хронического больного

Данные о каждом хроническом больном заносятся и в общую базу знаний системы и на основе этих данных формируются карты здоровья района, города, области, которые может увидеть на своем экране сотрудник департамента и ТФОМС.

2.3 Появление нового лекарства и вопросы фармэкономики

Рассмотрим ситуацию, когда менеджер фармакологического предприятия на экране своего АРМ добавляет в онтологию портала описание нового лекарства из некоторой группы препаратов.

Агент менеджера ФАРМБОКС, основываясь на базе знаний медицинского Интернет-портала, формирует список заинтересованных пользователей лекарства: менеджеров Департамента, всех главврачей и врачей, которые используют препараты данной группы для лечения своих пациентов. Менеджеру-фармацевту предлагается отправить сообщение о новом лекарстве с возможностью отредактировать список рассылки.

Агент менеджера ФАРМБОКС сообщает агенту главврача о новом лекарстве. Агент главврача на основании информации из базы знаний поликлиники выявляет пациентов, которым может быть назначено лекарство и рассчитывает возможную экономию от замены используемого препарата на новый. При этом учитывается не только сиюминутная экономия. Прогнозируется использование сэкономленных средств на льготные лекарства для профилактики отдельных пациентов и возможная экономия за счет снижения затрат на вызовы скорой помощи и госпитализации для этих пациентов. Более эффективное новое лекарство может быть даже

дороже существующих аналогов, но при этом его использование может дать существенную последующую экономию при сокращении продолжительности лечения, уменьшении затрат на вызовы скорой помощи и госпитализации, связанные с лечением основного заболевания. Все это учитывается при расчетах.

Агент главврача предлагает главврачу апробировать новое лекарство с указанием количества пациентов, которым оно может быть назначено, и возможную экономию, которую можно получить заменой ранее использованного препарата на новое лекарственное средство. При согласии главврача его агент сообщает агентам ВОП о принятом решении, а агенты ВОП формируют списки пациентов, которым может быть назначен препарат с указанием диагноза по каждому больному и возможной экономией. Эта информация доводится агентами до ВОП, им предлагается связаться с пациентами и назначить (рекомендовать) им новый препарат.

Врач при завершении случая по каждому больному отмечает эффективность лекарств, используемых при лечении. В этом ему помогает агент, выводя предварительный рейтинг лекарств на основании продолжительности срока лечения и его результатов.

По истечении срока апробации агент главврача собирает информацию о применении апробируемого лекарства, выявленном рейтинге лекарства и полученной экономии. На основании этой информации формируется предложение главврачу о необходимости послать в департамент запрос на включение нового препарата в список жизненно важных лекарственных средств (ЖВЛС). Пример такого сообщения представлен на рисунке 7. Главврач может выбрать продолжение апробации, отказ от использования неэффективного лекарства или посылку запроса в департамент о включении препарата в список ЖВЛС.

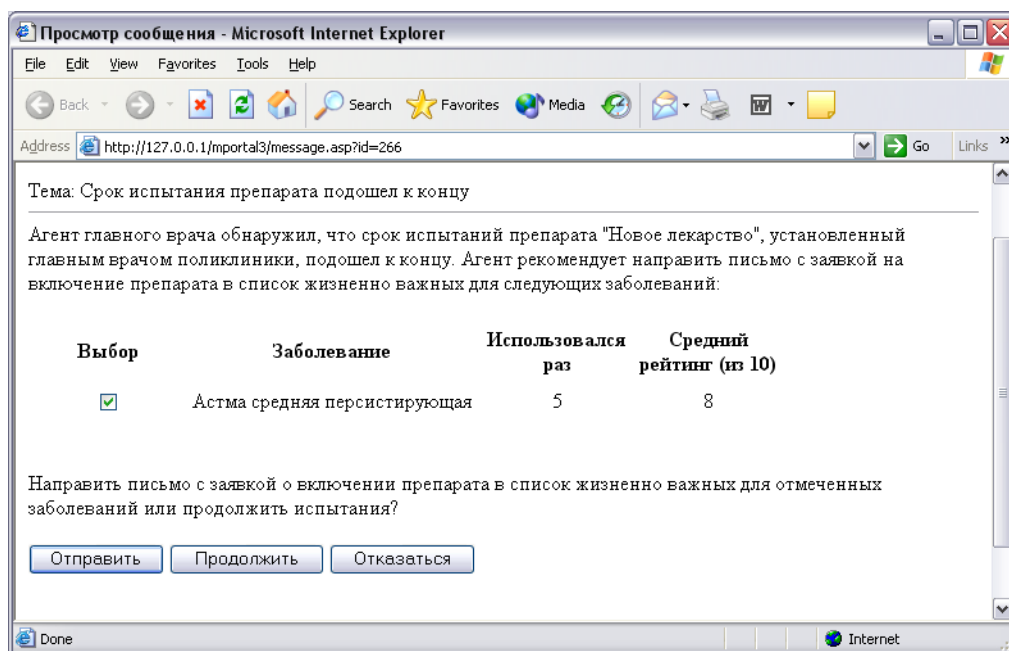


Рисунок 7 – Предложение агента главврача о включении нового лекарства в список ЖВЛС

Если главврач соглашается посылать запрос в Департамент, то агент главврача сообщает агенту менеджера Департамента о выявленном рейтинге, эффективности и полученной экономии по результатам применения лекарства. На основании настроек агент менеджера Департамента предлагает включить препарат в список ЖВЛС. В случае положительного решения агент менеджера сообщает о включении препарата в список ЖВЛС менеджеру ФАРМБОКС и глав-

врачам, и на основании информации от главврачей в дальнейшем планирует закупки этого препарата.

2.4 Постоянный мониторинг за состоянием пациента.

Интеграция агентов поликлиники, скорой помощи, стационара

В приемном покое стационара фиксируется факт поступления больного и его предварительный диагноз. Агент врача стационара или агент врача скорой помощи сообщает агенту ВОП поликлиники, наблюдающего данного больного, о факте поступления больного в стационар, его диагнозе и запрашивает его электронную карту.

Агент врача поликлиники формирует для ВОП сообщение о госпитализации и в зависимости от диагноза госпитализации и клинических протоколов указывает лечащему ВОП о рекомендуемых изменениях в схеме лечения основного заболевания пациента. Эта информация доводится до сведения ВОП и отправляется агенту врача стационара.

По завершении лечения в стационаре электронная карта с полной информацией о проведенном лечении и рекомендациями по последующему амбулаторному лечению и профилактике поступит в поликлинику по месту обслуживания пациента.

Схема взаимодействия агента врача стационара и агента ВОП поликлиники показана на рисунке 8.

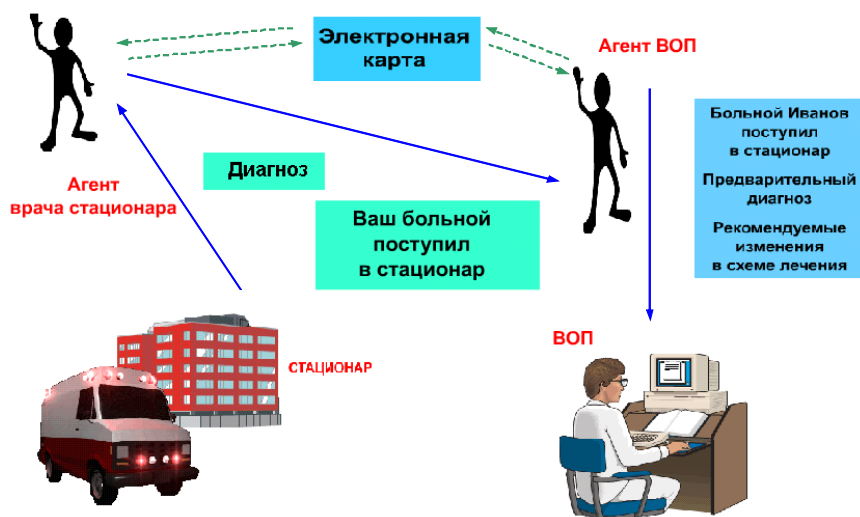


Рисунок 8 – Схема взаимодействия агента врача стационара и агента ВОП поликлиники

Приведенные примеры не исчерпывают ситуаций взаимодействия, воспроизводимых в модели. Помимо вышеописанных сценариев моделируются ситуации появления новой методики лечения и средств диагностики, роста количества заболевших (эпидемии), поиска нового врача, перепрофилирования дневного стационара, и в каждой такой ситуации при принятии решений используется мультиагентный подход и медицинские базы знаний.

3 Контроль показателей эффективности и качества работы системы здравоохранения и отдельных ВОП

Рассматриваемая сетевая мультиагентная модель системы управления здравоохранением была ориентирована не на “вертикальный” охват всего объема деятельности каждой организации, а на поддержание “горизонтальных связей” между ними, гарантирующее гибкость, оперативность и эффективность их взаимодействия. В дальнейшем это позволит по иному подойти к

описанию процессов в каждой организации, напрямую исходя из того, чем она полезна и эффективна для других организаций.

Тем не менее, уже в рассмотренной версии на автоматизированном рабочем месте менеджера Министерства здравоохранения региона моделируется мониторинг большинства показателей результативности деятельности Министерства и всей сферы здравоохранения в регионе. На рисунке 9 показана карта Самарской области, представленная на экране АРМ менеджера Министерства, где будут постоянно видны очаги проблем.

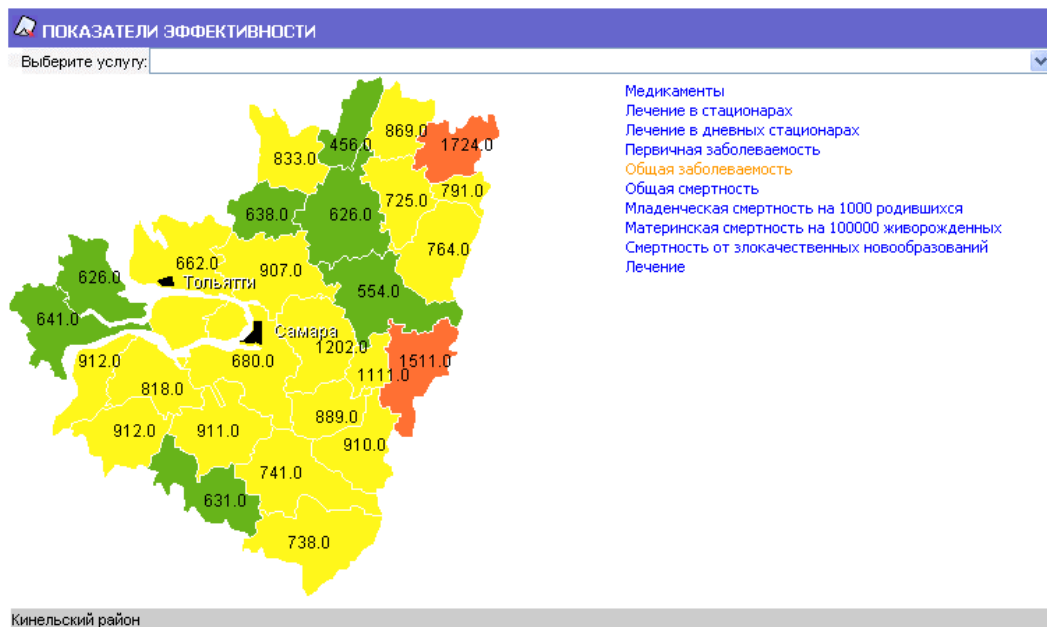


Рисунок 9 – Карта Самарской области на рабочем месте менеджера Министерства здравоохранения

Главный врач поликлиники на своем АРМ получил возможность осуществлять оперативный контроль за медико-экономическими показателями работы каждого врача, каждого участка и поликлиники в целом на основании информации из реальных баз данных поликлиники и экспертных оценок АО отдельным показателям. К таким показателям эффективности и качества работы ВОП, применяемой в ММУ Городская поликлиника № 15 относятся:

- количество обслуживаемых пациентов,
- заболеваемость на участке,
- количество присвоений категории инвалидности,
- количество реабилитированных пациентов,
- количество (стоимость) плановых и экстренных госпитализаций,
- количество (стоимость) вызовов скорой помощи,
- стоимость льготных медикаментов,
- процент вакцинации, флюорография,
- заполнение электронных амбулаторных карт,
- онкоосмотры, запущенные случаи онкологии,
- диспансеризация, профилактическая работа (школы, анкетирование).

Уже сейчас большинство этих показателей по отделениям и каждому врачу поликлиники отображаются на экране АРМ главврача в табличной форме и форме диаграмм изменения во времени по каждому выбранному показателю. Значения этих показателей по каждому участку

можно увидеть и на карте района обслуживания поликлиники, отображаемой на экране АРМ главврача и представленной на рисунке 10.

Система оперативного контроля медико-экономических показателей работы отдельных врачей, отделений и поликлиники в целом позволяет перейти к автоматизации процессов управления качеством [3].

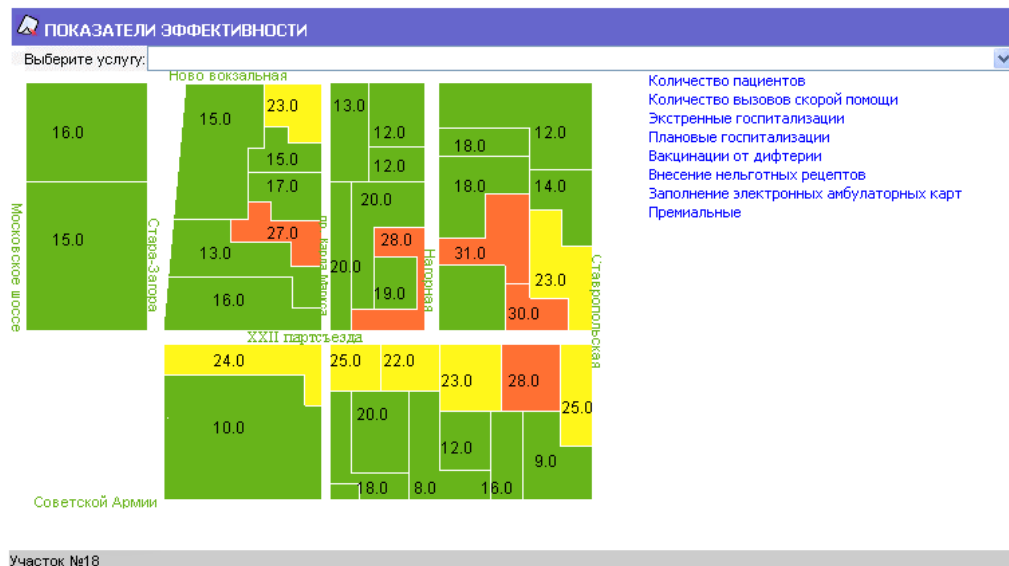


Рисунок 10 – Участки обслуживания поликлиники

Основные результаты рассмотренного проекта в 2004 году будут интегрированы в систему управления регионом и получат свое развитие в рамках проекта “Разработка первой очереди системы управления регионом с применением мультиагентных технологий” ФЦП “Электронная Россия”.

Список литературы

- [1] В.А. Виттих, Г.И. Гусарова, Г.П. Котельников, Ю.К. Ларионов, В.В. Павлов, П.О. Скобелев, Т.В. Тяпухина, Л.С. Федосеева, М.А. Шамашов, Е.В. Чернов. Базовые направления реформы здравоохранения в Самарской области и компьютерная модель для их апробации. // Труды 5-й Международной конференции по проблемам управления и моделирования в сложных системах. Самара, 17-22 июня 2003. – Самара: СНИЦ РАН, 2003, с. 414-423.
- [2] В.А. Виттих, В.Н. Ежков, Г.П. Котельников, Ю.К. Ларионов, В.В. Павлов, П.О. Скобелев, Л.С. Федосеева, М.А. Шамашов. Разработка интегрированной мультиагентной системы для управления здравоохранением в регионе. // Труды 4-й Международной конференции по проблемам управления и моделирования сложных систем, Самара, 17-24 июня 2002. - Самара: СНИЦ РАН, 2002, с. 398-406.
- [3] Л.С. Федосеева, Н.И. Кантемирова, И.М. Назаркина, Л.А. Щелкунова. Системный подход к управлению качеством оказания медицинской помощи в амбулаторно-поликлинических условиях. Наст. сборник.

МУЛЬТИАГЕНТНЫЙ ИНТЕРНЕТ-ПОРТАЛ ДЛЯ ИНТЕГРАЦИИ РЕСУРСОВ ДЕПАРТАМЕНТОВ СОЦИАЛЬНОГО БЛОКА САМАРСКОЙ ОБЛАСТИ

В.А. Виттих¹, П.О. Скобелев², О.И. Лахин², А.Н. Сальков²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61
cscmp@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

²ООО «НПК «Генезис знаний»
443011, Самара, ул. Ново-Садовая, 221, офис 206
lahin@kg.ru
тел./факс: +7 (8462) 17-30-69

Ключевые слова: социальная защита, здравоохранение, культура, модель интересов и предпочтений пользователя, возможности и потребности, открытые мультиагентные системы, онтология

Abstract

This article describes development of the multi-agent portal for integration of resources of the social block departments of the Samara Region. The portal provides integration of organizations of the social sphere of the region and provides interaction between them in order to satisfy the needs and abilities of the citizen of the region. The described Internet-portal is an open and developing system. The main distinctive feature of the system is using multi-agent approach, when each user (an organization or a citizen) and each resource (resources of organizations, needs and abilities) receive their own agent, which on the basis of ontological knowledge carries out constant monitoring of the recurrent content of the system and aims at satisfying the needs of its user.

Введение

Мультиагентный Интернет-портал для интеграции ресурсов департаментов социального блока Самарской области создан в рамках Первой очереди системы управления регионом с применением мультиагентных технологий и предназначен для повышения качества и эффективности управления в регионе, для реализации потребностей и возможностей граждан и организаций, а также обеспечения устойчивого социально-экономического развития региона.

На первом этапе портал должен решить следующие задачи: представлять лучшие достижения Самарской области в социальной сфере, здравоохранении и культуре; пополнять и развивать информационные ресурсы о деятельности людей и организаций, реализации проектов и происходящих событиях на территории Самарской области; описывать и формализовано представлять знания о структуре и функциях организаций; конструировать функциональные модели субъектов деятельности, описывать их возможности и потребности; способствовать развитию связей между людьми и организациями Самарской области и других регионов нашей страны; развивать индивидуальный адресный подход к каждому человеку, учитывающий его реальные потребности и возможности; повышать гибкость и эффективность деятельности государственных учреждений социального блока за счет учета обратной связи и конечного результата работы.

Интернет-портал позволяет менеджерам организаций социальной сферы создавать собственные «представительства» в системе и определять функции своих организаций и способы их

взаимодействия, как с другими организациями, так и с гражданами, через описание своих потребностей и возможностей.

На этой основе могут быть построены отраслевые подсистемы управления, включающие в качестве организаций для департамента социальной защиты - районные управления, социальные и реабилитационные центры, санатории, пансионаты, приюты и т.д.; для департамента здравоохранения - территориальные отделения фонда медицинского страхования (ТФОМС), уполномоченные региональные компании по снабжению льготными лекарствами (ФАРМ-БОКС), поликлиники, больницы, диагностические центры и т.д.; для департамента культуры - театры и музеи, библиотеки, клубы, культурные центры и т.д.

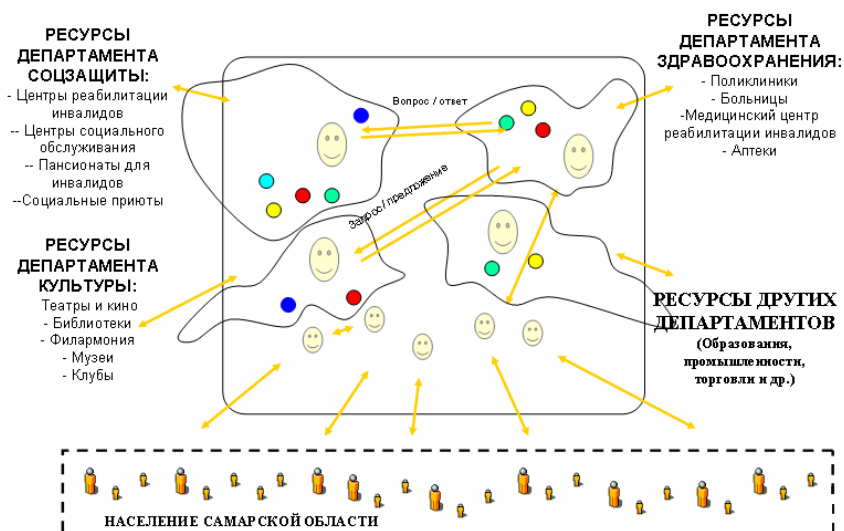


Рисунок 1 – Участники портала

При идентификации человека и с его разрешения, указанные организации получают возможность делать ему предложения и взаимодействовать с ним для удовлетворения его потребностей. Например, временный переселенец будет приглашен в конкретную поликлинику, больному инсулином после постановки диагноза сразу придет предложение центр «Диабет», на новую выставку икон древней Руси в художественном музее будут приглашены классы школ, изучающие по программе данную тему, поликлиника сможет напрямую взаимодействовать с ФОМС по вопросам инвестиций для приобретения оборудования и т.д.

В результате, портал обеспечивает интеграцию организаций социальной сферы региона и активизацию взаимодействия между ними для реализации потребностей и возможностей граждан региона.

1 Предлагаемый подход

Рассматриваемый Интернет-портал является *открытой и развивающейся системой*, что позволяет всем категориям пользователей не только получать самую интересную и полезную для них информацию, но и принять активное участие в повседневной жизни портала и развитии его ресурсов.

Главная особенность системы связана с применением *мультиагентного подхода* [1], при котором каждый пользователь (организация или гражданин) и каждый ресурс (ресурсы организации, возможности или потребности) получают своего агента, который на основе знаний из онтологии постоянно производит мониторинг периодически обновляемого контента системы и стремится удовлетворить интересы своего пользователя.

Под «агентом» следует понимать программный объект, способный воспринимать информацию из среды, планировать свои действия и принимать решения, действовать в среде и коммуницировать с подобными себе агентами. Агент действует на основе онтологии, в которой указаны как законы мира, так и возможные сценарии поведения агента. В результате знания предметной области агенты превращаются из пассивных и статичных компонент в «живые», активные и динамичные сущности, способные самостоятельно договариваться о взаимодействии, трансформироваться и устанавливать связи между собой.

Мультиагентные системы могут быть построены как открытые, способные «на лету» адаптироваться к изменениям за счет своих онтологий – баз знаний предметной области, отделенных от программного кода системы, пополнение которых открыто и доступно для специалистов и менеджеров, не имеющих навыков программирования. Эти системы способны накапливать, обрабатывать и применять знания о предметной области.

Применение онтологического подхода, играющего ключевую роль в способности агентов принимать решения, можно определить как совокупность пополняемых баз знаний, описывающих основные понятия и отношения предметной области, а также сценарии или правила поведения агентов в форме семантических сетей. Эта семантическая сеть может содержать как декларативные, так и процедурные компоненты и включать такие понятия (концепты) как физические или абстрактные объекты, свойства, процессы, отношения, атрибуты.

Каждому пользователю в разрабатываемом подходе дается программный агент, который действует от имени и по поручению человека для реализации его потребностей и возможностей. Такой программный агент, настраиваемый человеком на решение конкретных задач, может обнаруживать важные для него события, вырабатывать варианты решений и согласовывать принимаемые решения со всеми участниками взаимодействия. Например, для пенсионера, являющегося ветераном труда - обнаружить распространяющиеся на него дополнительные льготы, для переселенца-инвалида – определить размер материальной помощи, найти организацию, которая сможет предоставить ему возможность работы на дому, если такая потребность была заявлена человеком; больному диабетом – подобрать лечебное учреждение и т.д.

Для того, чтобы агенты могли находить варианты решений индивидуально для каждого человека, с учетом конкретных жизненных ситуаций, они получают доступ к персональным данным и базам знаний (онтологиям), которые могут грамотно интерпретировать и применять в интересах своего владельца. Описание онтологии при этом представляет собой семантическую сеть понятий и отношений, характерных для социальной сферы. Через семантическую сеть экземпляров этих понятий и отношений представляются конкретные ситуации каждого гражданина, к которым применяются хранящиеся в базе знаний правила рассуждений и принятия решений. В разрабатываемом подходе это позволяет строить динамические сети потребностей и возможностей (ПВ-сети) как для граждан, так и для организаций социальной сферы. Появление, например, нового закона, изменяющего социальный статус человека, или изменение его возможностей при этом сразу вызывает изменение в конфигурации ПВ-сети конкретного человека и с его согласия запускает цепочку взаимодействий и согласованного пересмотра и принятия решений во всех участвующих организациях.

Результаты взаимодействия граждан с учреждениями социальной сферы могут составлять важную исходную информацию для работы менеджеров по организации процессов управления департаментов социального блока, которые смогут оценивать качество и эффективность предоставляемых населению услуг, своевременно обнаруживать проблемные ситуации (нехватка финансовых средств на выдачу пособий, задержка с оформлением льгот и т.д.) и принимать решения по повышению эффективности используемых ресурсов.

2 Онтология ресурсов департаментов социального блока

В портале все ресурсы описываются на основе онтологии, которая специфицирует ключевые понятия и отношения (рисунок 2).

В структуре онтологии ресурсов департаментов социального блока Самарской области выделяются два уровня. Первый уровень (интенциональная сеть) - это уровень онтолог-разработчика системы, который создает сеть и следит за ее развитием по различным направлениям, иногда вводит новые и реструктурирует существующие базовые понятия. Второй уровень (экстенциональная сеть) – это уровень конечного пользователя, который пополняет как правило экземпляры различных классов семантической сети, описывая конкретный новый драматический театр, народный хор и т.д, а также экземпляры отношений между экземплярами различных классов [3].

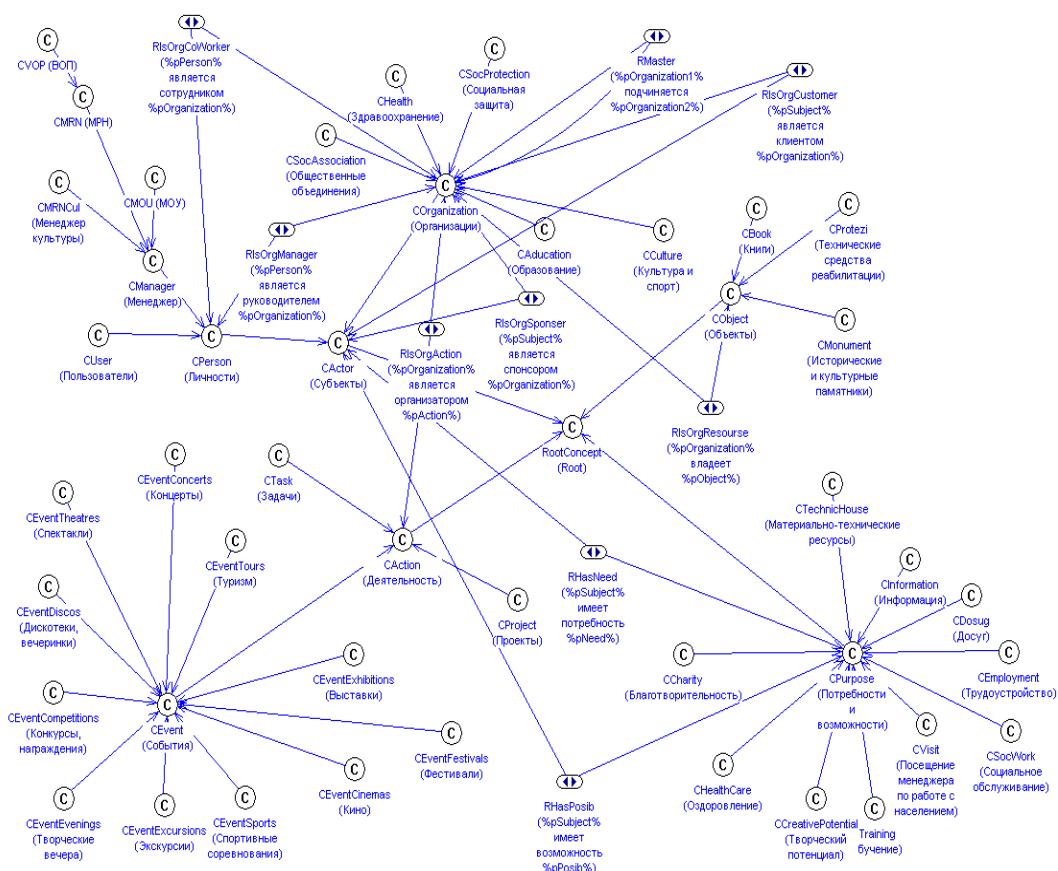


Рисунок 2 – Онтология ресурсов департаментов социального блока Самарской области

Онтолог-разработчик системы определяет базовые понятия (классы) системы, а также отношения, на основе которых прикладные пользователи будут описывать свои ресурсы. Эти ресурсы задаются путем инстанциаций объектов соответствующих классов, придания им некоторых свойств или атрибутов и связывания их определенными отношениями. Возможность динамического коллективного развития семантической сети ресурсов департаментов социального блока Самарской области – одна из ключевых особенностей разработанной системы [2].

2.1 Базовые классы онтологии

Все основные понятия, присутствующие в онтологии, описаны в виде классов. Классы онтологии системы описывают концепты из области взаимодействия организаций в социальной сфере. Классы делятся на две группы – базовые классы и динамические классы. Базовые классы жестко описаны в статической части онтологии, их структура неизменна, они не могут быть удалены. В то же время динамические классы являются наследниками базовых онтологических классов и описываются пользователями или инженерами по знанию по мере эксплуатации системы.

Базовыми классами онтологии являются: *Субъекты (Организации, Личности), Объекты, Деятельность (События, Проекты, Задачи), Возможности и потребности*. Класс *Субъект* – это базовый класс для описания различных субъектов деятельности (организации, физические лица). Класс *Объект* является базовым классом для описания всех материальных ресурсов. Класс *Деятельность* предназначен для описания организацией любого вида деятельности, а также задач, который она может выполнять. Класс *Возможности и потребности* предназначен для описания субъектом деятельности своих возможностей, что умеет делать и потребностей, в чем нуждается. Возможность для одной организации может являться потребностью другой организации.

2.2 Динамическая часть онтологии

Динамическая часть онтологии хранится в БД и содержит набор определенных пользователями классов, экземпляров, ограничений и отношений. Администратору системы предоставляется возможность самостоятельно описывать потомков базовых классов, а так же задавать правила, при которых устанавливаются отношения между экземплярами классов или самому устанавливать такие отношения.

Динамическими потомками базовых классов могут быть следующие понятия: у Организации - *Организации социальной защиты, Санатории, Пансионаты, Приюты, Организации здравоохранения, Поликлиники, Больницы, Диагностические центры, Аптеки, Организации культуры, Библиотеки, Театры, Кинотеатры, Музеи, Дома культуры, Творческие коллективы, ВУЗы, Школы, Детские сады* и т.п.; у Личности - *Руководители, Актеры, Преподаватели, Работники здравоохранения, Работники социальной сферы, Политики, Писатели, Художники, Ученые* и т.п.; у События - *Спектакли, Выставки, Конференции, Презентации, Визиты, Концерты, Кинофильмы, Конкурсы, Фестивали* и т.п.; у Возможности и Потребности - *Информация, Обучение, Творческий потенциал, Досуг, Благотворительность, Трудоустройство, Оздоровление, Социальное обслуживание, Материально-технические ресурсы* и т.п.

У каждого динамически описанного класса, могут быть динамически созданы атрибуты, характерные только этому классу.

2.3 Описание организации

При создании структуры организации, необходимо описать атрибуты, а также отношения, которыми организация связана с различными ресурсами. Базовый класс *Организация* имеет следующие атрибуты: *Наименование организации, Адрес организации, Контактный телефон организации, Направление работы*. Эти атрибуты являются статическими у базового класса организации. У динамических потомков класса организации, список атрибутов можно расширять, например, это могут быть: *Производственные площади, Вместимость зала, Наличие сцены, Наличие диагностического оборудования* и т.п.

Отношения, связывающие базовый класс *Организация* с другими базовыми классами: сотрудники организации (*Личность является сотрудником Организации, Личность является руководителем Организации*), материальные ресурсы организации (*Организация владеет Объект*), мероприятия проводимые организацией (*Организация является организатором Мероприятия*), клиенты организации (*Субъекты является клиентом Организации*), спонсоры организации (*Субъекты является спонсором Организации*), партнеры организации (*Субъекты*

является партнером Организации), возможности организации (Организация имеет возможность Возможность), потребности организации (Организация имеет потребность Потребность), подчиненные организации (Организации подчиняется Организация)

3 Использование портала

Для получения всех предоставляемых порталом сервисов пользователю необходимо зарегистрироваться в системе и описать карту интересов (рисунок 3), а также настроить свои возможности и потребности. При этом ему будет выделен персональный агент, который будет принимать участие в переговорах с агентами всех ресурсов, для выявления наиболее интересных.

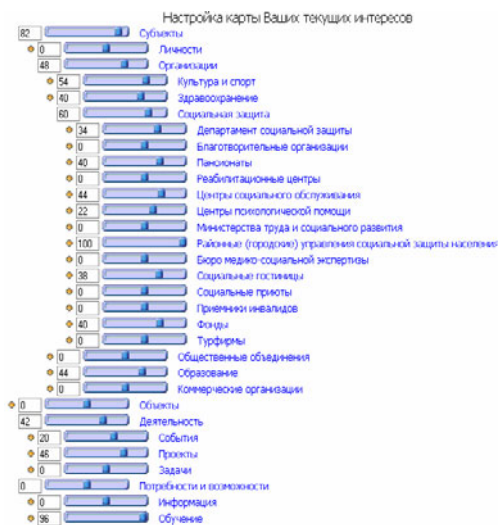


Рисунок 3 – Настройка карты интересов

ресурса, так как пользователь не указал явно, что найденные ресурсы ему интересны. Например, если пользователь указал уровень интереса к кинотеатру “Звезда” равными или выше 20%, то ему будет приходиться оповещение о событиях, связанных с этим кинотеатром.

Кроме явной настройки пользователем своей карты интересов в портале имеется возможность накопления неявной карты интересов пользователя. Динамическая модель неявных предпочтений пользователя меняется самой системой посредством сбора информации о пользователе (какие разделы посещает чаще, куда совсем не заходит), и дальнейшего ее анализа.

Пользователи и организации также могут настроить свои возможности и потребности (рисунок 4). В списке возможностей отображаются названия, описанных возможностей. Чтобы редактировать параметры возможности необходимо щелкнуть на названии возможности, после этого открывается форма для редактирования параметров возможности. Для удаления возможности необходимо щелкнуть на кнопке *Удалить* напротив названия возможности, при этом из системы удалится экземпляр возможности. Для добавления новой возможности необходимо щелкнуть на кнопку *Добавить*, при этом в рабочем поле появляется форма для описания новой возможности.

При описании новой возможности необходимо выбрать класс возможности из выпадающего списка. При выборе разных классов возможности меняются поля *Дополнительные настройки возможности* и *Возможные отношения*, которые зависят от онтологического описания разных классов. Далее необходимо описать Заголовок, Картинку и Текст рассылки с пред-

Карта интересов выводится в виде дерева классов понятий и экземпляров предметной онтологии, и пользователь может указать свою степень интереса к каждому из них. Настройка карты интересов (рисунок 3) связана с настройкой уровня слежения за типами (классами) ресурсов или определенными экземплярами ресурсов, а также параметрами извещения.

С помощью фильтра слежения можно настроить ограничения на оповещение по e-mail о появлении новых ресурсов или типов (классов) ресурсов в портале. Т.е. на e-mail пользователя будут посылаться извещения о появлении в системе только тех новых ресурсов или классов ресурсов, уровни интересов которых выше 0%. Более того, если пользователь укажет достаточно высокий уровень интереса к ресурсу, то система произведет поиск ресурсов, связанных с текущим через отношение, при этом их уровень интереса будет равен 70% от уровня интереса текущего

"Предоставление льготных путевок"

Свойства:

Текст рассылки:

Период рассылки:

Пол:

Возраст:

Социальный статус:

Стижмость:

Ресурсы возможности:

Социальные категории:

Целевая аудитория:

- Поиск предложений по типу возможности
- Разослать на выбранные социальные категории
- Разослать на граждан с выбранным социальным статусом
- Разослать на граждан, пользовавшихся услугами организации
- Разослать предложения на граждан с аналогичными свойствами
- Разослать предложения на граждан с аналогичной возможностью
- Разослать на широкую аудиторию

Рисунок 4 – Описание возможности

сообщение имеет дату, заголовок, источник сообщения, статус. При щелчке на заголовке сообщения осуществляется переход на субъект деятельности, от которого пришло сообщение или с которым оно связано. При щелчке на статус отображается реакция пользователя на сообщение, может быть: Отклонил, Принял, Не определено. Пользователи также имеют возможность оценить услугу и оставить свое мнение.

Все эти сообщения также дублируются на e-mail пользователей.

Портал обладает также рядом дополнительных сервисов таких как просмотр всех ресурсов портала, обсуждение ресурсов, выставление оценки и просмотр рейтинга ресурсов, просмотр списка событий (мероприятий), поиск по ключевым словам, общение пользователей в режиме форума и т.п.

Интернет-портал реализован на базе инструментальной системы MAWAE [1].

Заключение

В настоящей работе представлен мультиагентный Интернет-портал для интеграции ресурсов департаментов социального блока Самарской области, созданный в рамках Первой очереди системы управления регионом с применением мультиагентных технологий

В портале каждый ресурс и каждый пользователь получают агентов, которые действуют на основании своих стратегий, настраиваемых пользователем, получают возможность принимать решение и договариваться между собой о том, какой ресурс портала будет являться наиболее интересным и необходимым пользователю. Портал является открытой саморазвивающейся системой, содержание которой может пополняться всеми заинтересованными организациями и пользователями. Портал обеспечивает интеграцию организаций социальной сферы региона и активизацию взаимодействия между ними для реализации потребностей и возможностей граждан региона.

Таким образом, данная система обеспечивает сквозную, от уровня взаимодействия с гражданином – до уровня управления ресурсами департаментов социального блока, поддержку принятия решений, базирующуюся на знаниях, что позволяет менеджерам организаций самостоятельно пополнять онтологии и развивать систему и, тем самым, сократить затраты на разработку и сопровождение системы.

ложением по этой возможности. Этот текст будет рассылаться на целевую аудиторию. После этого можно настроить целевую аудиторию, на которую пойдет рассылка (рисунок 4). Помечая или отменяя галочку можно активизировать рассылку на заданную аудиторию. Описание организации и гражданами своих потребностей аналогично описанию возможностей.

После описания субъектами деятельности своих возможностей и потребностей, настройки целевой аудитории и периода активизации агента, запускается матчнинг на поиск возможных предложений. Активизация агента происходит периодически через указанный промежуток времени настраиваемый в периоде активизации агента.

Найденные предложения отображаются в виде списка сообщений и поступивших предложений, отсортированных по дате. Каждое сообщение

Список литературы

- [1] Батищев С.В., Лахин О.И., Минаков И.В., Ржевский Г.А., Скобелев П.О. Разработка и реализация пакета программ для построения мультиагентных приложений в сети Интернет // Труды Самарского научного центра РАН, 2001.
- [2] Батищев С.В., Искварина Т.В., Скобелев П.О. Построение развивающихся онтологий открытых мультиагентных систем в сети Интернет // Труды 5-ой Международной конференции по проблемам управления и моделирования в сложных систем, Самара, 17 - 22 июня, 2003.
- [3] Батищев С.В., Генералова Л.Д., Горбунова Т.Ф., Искварина Т.В., Куруленко Э.А., Лахин О.И., Скобелев П.О. Онтология социокультурных ресурсов Самарской области // Труды 5-ой Международной конференции по проблемам управления и моделирования в сложных систем, Самара, 17 - 22 июня, 2003.

СИСТЕМНАЯ ДИНАМИКА В УПРАВЛЕНИИ РЕГИОНАЛЬНЫМ РАЗВИТИЕМ (НА ПРИМЕРЕ МУРМАНСКОЙ ОБЛАСТИ)*

А.В. Горохов, В.А. Путилов, А.А. Шестаков

Институт информатики и математического моделирования технологических процессов
Кольского научного центра РАН
184209 Мурманская обл., г. Апатиты, ул. Ферсмана 24А, Россия
gorokhov@iimm.kolasc.net.ru
тел: +7 (81555) 7-40-50, факс: +7 (81555) 7-42-26

Ключевые слова: системная динамика, модель, регион, бескризисное развитие, сценарий, выбор, стратегия

Abstract

The study presents the approach to the economical development strategy in Murmansk region based on the synthesis and analysis of the system dynamic models. The method allows to estimate the worked out strategies as to the implementation “possibility” in socioeconomic system. By the “movement” trajectories analysis of the dynamic model within its internal conditions bounds.

Введение

В настоящее время существенно возросли требования к рациональному обоснованию управленческих решений, влияющих на различные аспекты функционирования социально-экономических систем и реализующих стратегии их бескризисного развития. Мощным толчком для исследований в этой области послужили социально-экономические трансформации, произошедшие в России в начале 1990-х годов. Быстрое разрушение старого уклада жизни привело к повышению неопределенности в решении задач управления социально-экономическими системами, особенно регионального уровня, так как регионы получили большую самостоятельность, не имея достаточного опыта управления даже в старых условиях. В северных регионах вследствие экономического спада возникла неустойчивая демографическая ситуация, опасная необратимыми негативными изменениями в экономике и других сферах жизни. В современной экономической теории нет единого мнения о причинах происшедшего спада и предпосылках возможного экономического роста [1-3]. Остро встала проблема оценки последствий принимаемых решений, в частности на региональном уровне. В связи с этим в последнее время региональной тематике посвящено большое количество исследований. Это создание моделей территориального социально-экономического развития регионов [1], разработка механизмов экологически устойчивого развития [2], исследование проблем федерализма и местного самоуправления [3] и др.

1 Системная динамика в моделировании социально-экономических систем

Экспериментальные воздействия на региональные социально-экономические системы по многим причинам (ограниченные временные рамки, опасность необратимых изменений, высокая стоимость и др.) обычно невозможны или нежелательны, поэтому основным методом изучения и прогнозирования поведения таких систем служит моделирование. Моделирование социально-экономических систем с достаточной для получения практически значимых результа-

* Работа выполнена при поддержке РФФИ (проект 02-07-90074).

тов многосторонностью заставляет рассматривать их как сложные динамические системы с множественными внешними и внутренними связями, учитывать разнообразные информационные, финансовые, материальные, энергетические потоки, предусматривать анализ последствий изменения их структуры, кризисных ситуаций и т.п.

Одной из важнейших особенностей социально-экономических систем является наличие большого числа обратных связей и сильное взаимное влияние между параметрами системы. То есть в системе практически отсутствуют независимые управляющие параметры. Даже внешние взаимодействия во многом зависят от состояния и результатов функционирования самой системы. С середины 1980-х годов школа И. Пригожина [4] развивает подход, согласно которому в развитии любой сложной системы чередуются периоды, в течение которых система ведет себя то как «в основном детерминированная», то как «в основном случайная». На принципиальную невозможность полноты информации о сложной системе в окрестности бифуркации указывал академик Н.Н. Моисеев [5]. Таким образом, основной проблемой моделирования сложных систем при недостаточной информации об их функционировании является обеспечение адекватности создаваемой модели объекту моделирования. Задача такого моделирования выходит за рамки формальных постановок и существенно ограничивает возможность применения математических методов описания поведения систем на базе статистического материала.

Поиск путей преодоления проблем в изучении социально-экономических систем, связанных с этими особенностями, привел к появлению в 1960-х годах специализированного метода имитационного моделирования – метода системной динамики. Данный метод, предложенный Джейм Форестером [6], зарекомендовал себя как эффективный подход к изучению поведения сложных трудно формализуемых систем. Системная динамика в настоящее время успешно развивается в разных странах. Метод системной динамики позволяет исследовать поведение сложных систем, опираясь на возможности компьютерного моделирования. В отличие от «традиционных» методов компьютерного моделирования системная динамика не требует построения математической модели исследуемого объекта в традиционной форме, а дает исследователю инструментарий для моделирования: компьютерные модели системных элементов и связей между ними. Это позволяет строить модели совершенно разной степени сложности и быстро получать результаты моделирования для разных параметров моделей. Подобно всем мощным средствам, существенно зависящим от искусства их применения, системная динамика способна дать как очень хорошие, так и очень плохие результаты. Поэтому для принятия решений на основании результатов динамического моделирования важно ясно представлять смысл вводимых допущений при синтезе динамической модели. Процесс синтеза моделей системной динамики основан на ментальных моделях: динамические модели создают коллективы экспертов, и синтез приемлемой для практического использования динамической модели может занимать до нескольких лет. Авторами разработан оригинальный метод концептуального синтеза динамических моделей сложных систем. Метод основан на использовании концептуальной модели как средства формализации экспертных знаний, что позволило автоматизировать наиболее трудоемкий этап динамического моделирования – синтез модели системной динамики [7].

2 Системно-динамическая модель Мурманской области

Разработка системно-динамических моделей социально-экономической системы региона необходима для исследования различных сценариев его развития с целью выработки стратегии, соответствующей бескризисному развитию. Эта задача имеет ряд специфических особенностей:

- исследуются сложные слабо формализованные ситуации, в которых невозможно применение аналитических методов;
- необходимо исследовать поведения систем в ситуациях, которые ранее не встречались;

- необходимо исследовать ситуации, наблюдение которых осложнено большой длительностью их развития.

Здесь очевидна неприемлемость аналитических и статистических методов моделирования, а также натурального эксперимента. Кроме того, последний может привести к необратимым негативным последствиям, чего также нельзя допустить. Единственно приемлемым методом исследования в данном случае остается динамическое моделирование.

Концептуальное моделирование Мурманской области с целью ее бескризисного развития показало [8], что определяющими факторами (точками чувствительности) в настоящих внешних условиях является экономический потенциал области (в первую очередь, добывающая и перерабатывающая промышленность) и ее население. Также важными факторами являются инфраструктура, телекоммуникации и экологическая обстановка, но определяющими рост они в настоящих условиях не являются. Разумеется, в других условиях эта классификация факторов будет выглядеть иначе. Одна из характерных особенностей сложных динамических систем – это то, что в процессе развития системы могут изменяться ее «точки чувствительности», то есть набор факторов, определяющих ее поведение, зависит от времени и от параметров состояния системы. Например, очевидно, что в условиях экологического кризиса состояние окружающей среды будет самым важным фактором, определяющим дальнейшее развитие системы, а в условиях нормального (сбалансированного) состояния окружающей среды этот фактор не будет оказывать практически никакого влияния на развитие системы.

Для развития экономического потенциала необходимы соответствующие условия. Определяющими для экономического потенциала Мурманской области являются экономические, финансовые и «материальные» условия, под которыми подразумевается технологический и технический уровень производственной базы области. Экономические условия области влияют на производство и финансовые условия через цены и налоги. Финансовые условия влияют на «материальные» условия через капиталовложения и инвестиции. «Материальные» условия влияют на производство через технологии. Производство в свою очередь влияет на финансовые условия области качеством, себестоимостью и объемом производимой продукции. Все это определяет уровень жизни населения области и, следовательно, миграционные процессы. Более подробно динамическая модель Мурманской области и ее реализация средствами системы динамического моделирования Powersim рассмотрены в работе [9].

3 Выбор стратегий на основе системно-динамического моделирования

Задачи синтеза и рационального отбора стратегий регионального развития относятся к классу задач принятия решений при многих критериях. В этих задачах к решению предъявляется несколько требований и часть информации, необходимой для полного и однозначного определения требований к решению, отсутствует [10].

В терминах динамической модели социально-экономической системы регионального уровня под стратегиями развития понимаются совокупности темпоральных рядов взаимосвязанных значений управляемых параметров модели, обуславливающих требуемые траектории «движения» модели в пространстве ее внутренних состояний. Поэтому задача выбора стратегий может быть сформулирована следующим образом.

Множество A всевозможных временных рядов, приводящих к достижению цели на рассматриваемом временном интервале, считается множеством всевозможных стратегий (множество вариантов). Подмножество X множества A , которое удовлетворяет заранее заданным предметным (т.е. социально-экономическим) ограничениям будем называть множеством допустимых стратегий или предъявлением. Аналогично, множество $Y \subseteq X$ социально и экономически рациональных стратегий в данных условиях можно считать выбором. При этом формирование управленческих решений сводится к выбору элемента (или элементов) из множеств

ва Y . Поэтому процедуру формирования рациональных управленческих решений можно отождествлять с актом выбора из множества предъявления X его специального подмножества Y .

Формирование социально-экономических решений на региональном уровне осуществляется в условиях неполной информации о текущих ситуациях, что приводит к нечеткой интервальной (I -нечеткой) интерпретации подмножеств X и Y , заданных на универсуме A . Оперирование с I -нечеткими подмножествами требует корректных, математически обоснованных моделей нечеткого интервального выбора. Такие модели выбора могут быть формализованы в введенных теоретико-множественных терминах: множества вариантов (универсума) A ; I -нечеткого предъявления \bar{X} и I -нечеткого выбора \bar{Y} [11].

Возможны, по крайней мере, два типа моделей нечеткого интервального выбора: функциональные и процедурные. Для моделей функционального типа с конечными множествами вариантов A всегда существуют эквивалентные аналитические представления в виде соответствующих систем функций алгебры нечеткой интервальной логики [12]. Для процедурных моделей дополнительно требуется ввести отношение предпочтения (отношение порядка) степеней I -нечеткой принадлежности, которое может совпадать с сужением «естественного» частичного порядка на рассматриваемое подмножество нечетких интервальных константах.

Проблема нечеткого интервального выбора с помощью метода сечений может быть сведена к совокупности взаимосвязанных задач классической теории выбора вариантов. На рисунке 1 модель выбора представлена в следующем виде:

$$\bar{Y} = \frac{Ch : \wp(A) \setminus \emptyset \rightarrow \wp(A)}{\langle \sigma, \pi \rangle}$$

здесь Ch – функция выбора вариантов, $\wp(A)$ – булеан множества допустимых стратегий, а $\langle \sigma, \pi \rangle$ – механизм выбора, состоящий из нечеткого бинарного отношения на множестве стратегий (σ) и правила выбора (π).

Уровни сечения ($\alpha, \beta, \dots, \xi$) характеризуют достоверность выбора с соответствующими оценками интервала истинности (t, f) и степенью риска (R_S).

При таком подходе сохраняется возможность одновременно иметь математически обоснованные нечеткую надежность рекомендуемых рациональных управленческих решений и количественные оценки их модельной неточности.

Проведенный анализ дискриптивных возможностей нечеткого интервального выбора вариантов показывает что, для конечных и обозримых нечетких интервальных множеств предъявления (с точностью до лидирующей константы) более перспективной оказывается процедурная модель выбора вариантов управленческих решений, независимо от типа используемого отношения предпочтения.

Заключение

В работе предложен подход к разработке стратегии экономического развития Мурманской области, опирающийся на синтез и анализ моделей системной динамики. Такой подход позволяет оценивать разрабатываемые стратегии с точки зрения «допустимости» их реализации для социально-экономической системы путем исследования траекторий «движения» динамической модели данной системы в пространстве ее внутренних состояний. Предложена модель выбора стратегий развития региональных социально-экономических систем. Модель реализована с помощью аппарата интервальной нечеткой логики.

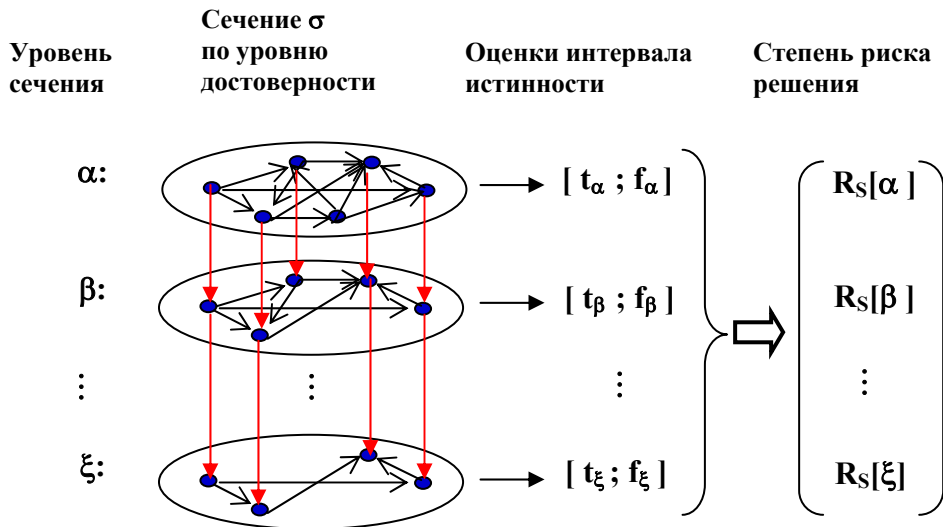


Рисунок 1 – Модель выбора стратегий.

Список литературы

- [1] Гранберг А.Г. Стратегия территориального социально-экономического развития России: от идеи к реализации // Вопросы экономики. 2001. № 9. С. 15-27.
- [2] Пегов С.А., Хомяков П.М. Моделирование развития экологических систем. - Л.: Гидрометеиздат, 1991. - 224 с.
- [3] Лексин В.Н., Швецов А.Н. Региональная политика России: концепции проблемы, решения // Российский экономический журнал. 1994. №2, №3; 1995. №3, №5-6.
- [4] Николис Г., Пригожин И. Познание сложного. Введение: Пер. с англ. – М.: Мир, 1990. 314 с.
- [5] Моисеев Н.Н. Математические задачи системного анализа. - М.: Наука, 1981. 488 с.
- [6] Forrester J. W. Industrial Dynamics - Portland, OR: Productivity Press, 1961. 464 pp.
- [7] Путилов В.А., Горохов А.В. Динамическое моделирование регионального развития // Информационные ресурсы России. 2003. № 4. С. 9-12.
- [8] Путилов В.А., Горохов А.В. Сложные динамические системы и моделирование // Информационные технологии в региональном развитии: концептуальные аспекты и модели. – Апатиты: Изд-во КНЦ РАН, 2002. С.6-16.
- [9] Путилов В.А., Горохов А.В. Системная динамика регионального развития. – Мурманск: НИЦ «Пазори», 2002. – 306 с.
- [10] Шестаков А.А. Логическое моделирование в условиях неопределенности. – Апатиты: Изд-во КНЦ РАН, 1996. 182 с.
- [11] Шестаков А.А. Интервальные обобщения теории нечетких множеств. – Апатиты: Изд-во КНЦ РАН, 1998. 40 с.
- [12] Шестаков А.А. Алгебра нечеткой интервальной логики. – Апатиты: Изд-во КНЦ РАН, 1998. 50 с.

СИСТЕМНЫЙ ПОДХОД К УПРАВЛЕНИЮ КАЧЕСТВОМ ОКАЗАНИЯ МЕДИЦИНСКОЙ ПОМОЩИ В АМБУЛАТОРНО-ПОЛИКЛИНИЧЕСКИХ УСЛОВИЯХ

Л.С. Федосеева, Н.И. Кантемирова, И.М. Назаркина, Л.А. Щелкунова

Муниципальное медицинское учреждение городская поликлиника №15
443111, Самара, ул. Фадеева, 56а
polik15@saminfo.ru
тел: +7 (8462) 51-97-35

***Ключевые слова:** системный подход, управление качеством, информационно-аналитическая сеть, медико-экономическая эффективность, мониторинг, стандарты, амбулаторный этап*

Abstract

In clause advantages of the system approach to quality management of rendering of medical aid at an out-patient-polyclinic stage are considered. For an example tactics of treatment of patients with a bronchial asthma, hypertonic illness and an anemia of pregnant women is chosen. The analysis and results of activity of doctors on achievement of clinical effect with various financial expenses is submitted to attention of readers.

Введение

Управленческий учет – это один из видов деятельности в системе управления. Его целью является обеспечение управленческого звена организации информацией оперативного и финансового учета, необходимой для анализа результатов финансово-хозяйственной деятельности организации, принятия на этой основе оптимальных управленческих решений и контроля за их реализацией. С развитием технических средств и совершенствованием информационных технологий стало возможным, с одной стороны, накапливать базы данных внутриучрежденческой информации, а с другой – получать незамедлительно всю необходимую для оперативной деятельности информацию и использовать ее в процессе управления. Если раньше внутриучрежденческая система информации ориентировалась в основном на внешних пользователей, то в современных условиях она выполняет прежде всего задачи обеспечения эффективного функционирования ЛПУ.

В ММУ ГП № 15 Промышленного района г. Самары создана локальная информационно-аналитическая компьютерная сеть, объединившая 130 рабочих мест. Информационные ресурсы формируются с 1996 года и включают в себя следующие базы данных:

- Медико-статистическая персонифицированная база данных оказанных услуг. Содержит информацию о медицинских услугах врачей и медсестер всех специальностей, параклинических подразделений и отделений функциональной диагностики, услуг дневных стационаров, стационаров на дому и неотложной помощи;
- Финансовая база данных. Содержит электронные реестры счетов за услуги, оказанные в других ЛПУ, результаты внутриучрежденческих взаиморасчетов;
- Данные о персонифицированном учете льготного лекарственного обеспечения;
- База данных застрахованного прикрепленного населения;
- Статистические сведения о состоянии здоровья населения, а также сведения о пациентах, имеющих факторы риска по развитию хронических заболеваний;
- Оперативные данные по кадровому составу ЛПУ;
- Банк нормативно-справочной информации, регламентирующей деятельность ЛПУ.

Сегодня информационные данные рассматриваются нами как инструмент для совершенствования управления не только ресурсами, но и качеством оказания медицинской помощи, что является одним из приоритетных направлений в период организации оказания первичной медико-санитарной помощи по принципу врача общей практики.

Основные составляющие подхода к управлению качеством медицинской помощи

Формирование института ВОП проходит в условиях явно недостаточного финансирования. Мощные социально-экономические реформы, проходящие в нашей стране, находят свое отражение в изменении существующих принципов организации здравоохранения. Одним из главных путей развития здравоохранения становится повышение его медико-экономической эффективности, предполагающей максимальную результативность медицинского учреждения, выраженную в улучшении качества оказания медицинской помощи, её доступности, а в дальнейшем в улучшении состояния здоровья населения.

В течение 8 лет на базе многопрофильной поликлиники ММУ ГП № 15 г. Самары проводятся реорганизационные мероприятия по оказанию медицинской помощи по принципу врача общей практики. Принятие ряда управленческих решений по организационно-структурным изменениям, непрерывной профессиональной подготовке врачебно-сестринского персонала, внедрение стационарозамещающих форм оказания медицинской помощи, развертывание «неотложки», отработка экономической модели деятельности ВОП, позволяющей создавать принципы экономической мотивации и осуществлять дифференцированную оплату труда, позволили на начальном этапе реформ стабилизировать экономическую ситуацию в поликлинике.

Однако опыт показывает, что адаптация системы оказания медицинской помощи по принципу ВОП к сложившимся жестким экономическим условиям должна проходить также с использованием эффективных механизмов управления качеством оказания медицинской помощи, т.к. ухудшение состояния здоровья населения в определенном проценте обусловлено и несовершенством оказания медицинской помощи. Международным сообществом разработаны основные принципы достижения лучшего качества: системный подход, использование научной методологии, работа в команде.

Использование современной методологии в работе ВОП поликлиники было начато на примере ведения больных, страдающих бронхиальной астмой, в 1996 г. Конкретным механизмом совершенствования медицинской помощи больным БА явилось внедрение в работу ВОП «формулярной системы» под редакцией академика А.Г. Чучалина, в основу которой положены принципы доказательной медицины, изложенные в материалах «Глобальная стратегия. Бронхиальная астма» (GINA). Работа проводилась при поддержке Департамента здравоохранения Администрации области и Управления здравоохранения г. Самары при непосредственном участии д.м.н., профессора Ю.Н. Осипова. На базе поликлиники был открыт пульмонологический кабинет, ВОП и медицинские сестры обучены новой методике диагностики, лечения, проведению мониторинга состояния больных, разработаны анкеты для скрининг-диагностики БА, контрольные карты для динамического наблюдения за пациентами, налажено снабжение медикаментами в соответствии с формуляром, открыта «Астма-школа». Были разработаны организационные и клинические алгоритмы на каждом этапе оказания медицинской помощи (пульмонолог, ВОП, МОП).

С целью оценки эффективности деятельности системы оказания медицинской помощи, больным бронхиальной астмой, своевременного «предотвращения отклонений» в ней, регулярно проводится оценка индикаторов качества: показатели заболеваемости, смертности, выхода на инвалидность, уровни вызовов скорой помощи и госпитализации, изучение качества жизни. Фармако-экономический анализ стал неотъемлемой частью работы ВОП.

Функционирующая в поликлинике локальная компьютерная сеть, позволяет персонифицировать учет событий, происходящих на всех уровнях системы. Созданы реестры пациентов,

имеющих право на льготное медикаментозное обеспечение, база данных всех услуг, оказанных больным внутри поликлиники и в других лечебных учреждениях города. Автоматизированная система обеспечивает анализ деятельности отдельных врачей в реальном времени и за выбранный календарный промежуток. Возможна оценка эффективности проводимой терапии препаратами различных групп с учетом изменения индикаторов эффективности и оценкой стоимости заболевания. Выход в сеть Интернет с 20 рабочих мест врачей позволяет активнее внедрять в практику работы врача современные достижения.

Системный подход в организации наблюдения за больными бронхиальной астмой принципиально изменил ситуацию, сложившуюся до 1995 года. Активное выявление больных с использованием скрининг-методов позволило увеличить количество наблюдаемых больных со 196 человек в 1996 г. до 1125 в 2003 г. Неоспоримым фактом является то, что адекватная базисная противовоспалительная терапия способствует контролю над состоянием больного. В 1995 году она сводилась в основном к назначению β_2 -агонистов короткого действия, системных кортикостероидов и ксантиновых производных. Контроль над симптомами достигался только у 12% больных. Показатель госпитализации превышал 35%. Использование современного «ступенчатого подхода», основанного на трактовке бронхиальной астмы как хронического воспалительного заболевания аллергического характера, позволило значительно изменить ситуацию.

За период работы с 1996 г. по 2003 г. показатель госпитализации с 35% снизился до 3%, а показатель вызовов скорой помощи со 145% до 21,1%. За прошедшие годы не зарегистрировано ни одного случая смерти от бронхиальной астмы, госпитализации в состоянии астматического статуса, в 3 раза снизился выход на инвалидность. В 5,8 раз уменьшились финансовые затраты на госпитализацию и вызовы скорой помощи. Прямые затраты на одного больного в год, по сравнению с другими лечебно-профилактическими учреждениями, почти в три раза меньше.

Убедительно подтверждают преимущества лечебной тактики, проводимой в поликлинике, изучение показателей «Качества жизни». Для исследования использовался стандартизованный международный опросник SF-36. Выявлено, что больные, находящиеся под наблюдением в поликлинике, имеют достаточно более высокий уровень качества жизни по следующим критериям: физическая активность, боль, общее здоровье, социальная активность, сравнение самочувствия с прошлым годом, жизнеспособность.

С целью совершенствования специализированной медицинской помощи больным бронхиальной астмой города Самары на базе пульмонологического кабинета ММУ ГП № 15 Приказом Управления здравоохранения Администрации г. Самары в 1998 году открыт городской амбулаторный «Астма-центр». Это смелое решение позволило современные принципы лечения больных бронхиальной астмой внедрить в практику всех районов города. Поддержка Департамента здравоохранения способствует распространению данного опыта в другие города и районы Самарской области. В штатное расписание лечебных учреждений введены кадры пульмонологов, открыты обучающие школы для пациентов. На базе «Астма-центра» проводятся конференции, практические семинары, тренинги, «круглые столы» по новейшим достижениям в лечении больных бронхиальной астмой с привлечением главных специалистов области и города, а также профессорско-преподавательского состава Самарского Государственного медицинского университета. За два года в «Астма-центре» проконсультировано более двух тысяч пациентов из других лечебно-профилактических учреждений города. Совместно с Центром профилактики Управления здравоохранения совершенствуется методическое пособие для врачей и медсестер, проводящих занятия в «Астма-школе».

Такой комплексный подход к организации пульмонологической помощи больным бронхиальной астмой позволил изменить ситуацию в городе в положительную сторону. На 52% увеличилось количество больных бронхиальной астмой, находящихся на диспансерном учете.

Хорошо эта работа поставлена в Железнодорожном, Промышленном, Советском и Красноглинском районах.

Наличие в Управлении здравоохранения Администрации г. Самары централизованной базы данных льготных рецептов на все медикаменты, назначаемые пациентам, страдающим бронхиальной астмой, позволил проанализировать эффективность проводимой терапии в районах города.

Проведенный анализ показывает, что увеличение процента всех финансовых ресурсов, затрачиваемых на приобретение базисных препаратов, приводит к снижению показателей госпитализации (индикатора качества оказания медицинской помощи). Так в Железнодорожном районе 24,5% всех средств затрачивается на приобретение ИКС, показатель госпитализации - 10%; в Промышленном районе - 28,9%, показатель госпитализации - 9,7%; в Кировском районе - 20,8%, показатель госпитализации - 15,6%. Расходование в ММУ ГП № 15 более 40% всех средств на препараты базисной терапии, позволяет поддерживать показатель госпитализации на уровне 3-4% в течение нескольких лет. Такая же зависимость выявлена при анализе вызовов СП. В этих же районах прослеживается тенденция снижения финансовых затрат на приобретение системных кортикостероидов, не являющихся базисными препаратами.

Медико-экономическая эффективность – это достижение клинического эффекта с наименьшими финансовыми затратами. В формулярах лекарственных средств всех районов присутствуют как дорогостоящие медикаменты, так и препараты менее дорогие. Можно отметить, что в ряде районов наибольшее предпочтение отдается дорогостоящим препаратам (Красноглинский, Ленинский, Октябрьский, Самарский). На флексотид и ингакорт затраты в рублях в год в Красноглинском районе составляют соответственно 143,5 рубля и 232,4 рубля; в Самарском районе – 172,1 и 168,9 рубля. В ММУ ГП №15 на эти дорогостоящие препараты затрачивается 16,2 рубля на флексотид и 56,2 рубля в год на ингакорт на одного больного в год. В то же время на отечественный ингаляционный кортикостероид, стоимость которого почти в 3 раза дешевле ингакорта и флексоотида, в ММУ ГП № 15 расходуется на одного больного в год 102,8 рубля, в Красноглинском районе он используется в 4 раза реже, а в Самарском не применяется вовсе. Это естественно не может не отразиться на «стоимости» бронхиальной астмы.

Показателями экономической эффективности лечения больных бронхиальной астмой могут служить «прямые» затраты на одного больного в год (это сумма затрат на приобретение льготных медикаментов, госпитализацию и вызовы скорой помощи). «Прямые» финансовые затраты в год на одного больного бронхиальной астмой в ММУ ГП № 15 по результатам 2002 года составляют 1061,5 рубля, а в Самарском районе – 4097 рубля.

Как положительный результат организации работы по ведению больных бронхиальной астмой на амбулаторном этапе необходимо отметить Железнодорожный, Промышленный, Ленинский, Кировский, Советский районы. Колебания «прямых» затрат в них составляют от 1725,2 рубля в год на одного больного в Железнодорожном до 2340 рублей в Кировском.

Таким образом, применение комплекса апробированных терапевтических методов, использование общеобразовательных школ для медицинских работников и пациентов, тщательный фармако-экономический анализ позволяют перевести лечение больных, страдающих бронхиальной астмой, на амбулаторно-поликлинический этап с улучшением качества оказания медицинской помощи и значительной экономией финансовых ресурсов.

Внедрение в практику новой медицинской технологии ведения больных бронхиальной астмой в ММУ ГП № 15, позволяет экономить более 1,5 млн рублей в год. Согласно разработанному «Положению оплаты труда ВОП», часть сэкономленных средств с учетом вклада каждого врача в использовании данной технологии, выдается как дополнительные средства к заработной плате, т.е. создается экономическая мотивация. Вторая половина сэкономленных ресурсов используется для улучшения лечебно-диагностической базы.

Учитывая положительный результат внедрения системного подхода в организацию работы ВОП по улучшению качества оказания медицинской помощи больным бронхиальной аст-

мой, в поликлинике создано 11 групп по адаптации существующих федеральных и разработке новых клинических протоколов. Их тематика актуальна в повседневной работе ВОП: хронические гепатиты, ИБС, сердечная недостаточность, заболевания суставов и т.д.

Работа по организации системного подхода в ведении больных, страдающих гипертонической болезнью началась в 1998 году под руководством д.м.н. профессора Н.Н. Крюкова. Проведенные скрининговые исследования по активному выявлению больных, верификации диагнозов позволили увеличить число наблюдаемых больных, страдающих гипертонической болезнью, с 1207 до 7529 человек, а также сформировать группу риска в количестве 4328 пациентов. Произошли изменения в назначении медикаментозной терапии. Сегодня предпочтение отдается препаратам, обладающим не только гипотензивным, но и цитопротекторным действием на органы-мишени. Значительно возросло применение пролонгированных ингибиторов АПФ, блокаторов каналов кальция, селективных β -блокаторов, тиазидоподобных диуретиков. Одновременно с этим значительно снизилось назначение медикаментов короткого действия.

Следующий шаг – организация обучающих школ для пациентов. Обеспечение больного информацией об особенностях течения заболевания, методах самоконтроля и самопомощи при развитии гипертонических кризов, позволяет избежать развития осложнений, улучшить качество медикаментозной терапии и уменьшить финансовые затраты. Проведенный медико-экономический анализ среди 45 пациентов, прошедших обучение в школе показал, что затраты на госпитализацию и вызовы скорой помощи снизились в 1,8 раза.

Наличие информационно-аналитической компьютерной базы данных состояния здоровья населения, оказанных медицинских услуг, лекарственного обеспечения, позволило провести анализ эффективности проводимых организационных мероприятий на группе больных в количестве 3734 человек, имеющих право на льготное лекарственное обеспечение. По сравнению с 1999 годом в 2002 году число вызовов скорой помощи снизилось на 35,4%, госпитализаций – на 14,5%, количество осложнений – на 11,4%, несмотря на то, что у этих больных кроме гипертонической болезни имеются заболевания, отягощающие течение гипертонии, увеличивающие риск сосудистых осложнений.

Системный подход использован нами при организации наблюдения за больными, страдающими хроническими обструктивными болезнями легких. И этот опыт подтверждает значимость назначения медикаментозной терапии, основанной на принципах доказательной медицины. Результаты ретроспективного анализа лечения больных ХОБЛ в 1999 году показали: ситуационное лечение больных с использованием бронхолитиков короткого действия, применение в 27% случаев необоснованной антибактериальной терапии, приводили к частым госпитализациям (30%), вызовам скорой помощи (103,7%), прогрессированию заболевания, что подтверждено результатами спирографического исследования.

Назначение базисной терапии согласно рекомендациям ведущих специалистов, изложенных в Глобальной стратегии по ведению больных ХОБЛ (QOLD 1998) и Федеральной программе 1999 года, изменили ситуации в положительную сторону. В качестве препаратов базисной терапии назначались бронхолитики различных фармакологических групп: М-холинолитики, β_2 -агонисты, их комбинации, метилксантины, а при их неэффективности – ингаляционные или системные глюкокортикостероиды. При таком подходе к лечению, в 2003 году у больных данной группы показатель госпитализации снизился в 8,1 раза, а показатель вызовов СП – в 4,6 раза. Уменьшились и финансовые затраты на госпитализацию и вызовы скорой помощи с 253,6 тыс. рублей в 1999 г. до 39,9 тыс. рублей в 2003 г.

То есть происходит перераспределение финансовых ресурсов с уровня стационара и скорой помощи на амбулаторно-поликлинический этап, что создает условия для увеличения объема оказания медицинской помощи населению, улучшения лечебно-диагностической базы поликлиники, повышению профессиональных знаний врачей и медицинских сестер, а значит и качеству оказания медицинской помощи. Сегодня в поликлинике созданы возможности для

лечения в спелеокамере, проведения небулайзерной терапии и кислородотерапии как курсами, так и длительно на дому.

Проблема охраны материнства и детства остается приоритетной на всех этапах оказания медицинской помощи. Наличие у женщин фертильного возраста в 84,6% экстрагенитальной патологии, большое количество осложнений во время беременности подтолкнули нас именно с позиции системного анализа подойти к организации оказания помощи беременным женщинам на амбулаторном этапе. В 1998 году 34,0% в структуре заболеваний беременных составляла анемия. До конца 2002 года этот показатель неумолимо возрастает до 44%, и параллельно отмечается рост осложнений беременности. Активное лечение анемии не только при явной клинической картине, а превентивная терапия с учетом прогнозирования вероятности развития, привели к положительным результатам. Процент регистрируемой анемии снизился до 35%, а количество осложнений на 13%. Решение данной проблемы потребовало не только подготовки врачей общей практики, гинекологов, акушерок, но и большой разъяснительной работы среди беременных женщин в обучающих школах и на индивидуальных приемах. Эта работа проводилась при поддержке главных специалистов Управления и Департамента здравоохранения, а также кафедры гинекологии, возглавляемой О.И. Линевой.

Заключение

Таким образом, применение комплекса современных апробированных терапевтических методов, использование образовательных школ как для медицинских работников, так и для пациентов, тщательный фармако-экономический анализ позволяет перевести лечение больных, страдающих хроническими заболеваниями на амбулаторно-поликлинический этап с улучшением качества оказания медицинской помощи и значительной экономией финансовых ресурсов.

Список литературы

- [1] The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. *Lancet* 1998;351:1225-32.
- [2] Кучеренко В.З. Некоторые подходы к оценке качества лечебно-диагностического процесса // Проблемы управления здравоохранением. 2003. №1(8). С.13-17.
- [3] Алмазов В.А. и др. Качество медицинской помощи и обоснования нового подхода его оценки в условиях развития системы медицинского страхования // В кн.: Пациенты и врачи за качество медицинской помощи. – СПб, 1994. С.35-62.
- [4] Оценка качества и эффективности медицинской помощи / Под ред. О.П.Щепина. – М., 1992.
- [5] Климова Н.Б. Развитие стандартизации в системе обязательного медицинского страхования: новый этап // Проблемы стандартизации в здравоохранении: Материалы IV научно-практической конференции (Москва, 27-29 марта 2002 г.). – Москва, 2002. С.10-14.
- [6] Вялков А.И. Управление качеством в здравоохранении // Проблемы стандартизации в здравоохранении: Материалы V научно-практической конференции (Москва, 24-26 марта 2003 г.). – Москва, 2003. С.3-11.

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ИНВЕСТИЦИЙ В НЕФТЕДОБЫЧЕ НА ОСНОВЕ МОДЕЛИРОВАНИЯ

Е.Н. Коновалов¹, В.П. Офицеров¹, С.В.Смирнов²

¹АКГ «Развитие бизнес систем»

113093, Москва, ул. Люсиновская, 27, стр. 3

E-mail: OfitserovVP@rbsys.ru

тел: +7 (095) 967-6838 факс: +7 (095) 967-6843

²Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

smirnov@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: инвестиции, нефтедобыча, моделирование, оптимизация, имитация

Abstract

The problem of planning of investments is considered by the example of the oil-extracting company. For effective decision of this problem the synthesis of optimization and simulation dynamic modeling is offered. The optimization model has universal character and is based on ideas of the construction of Bellman's function. The dynamic model is described conceptually and should be developed for each concrete case. Both approaches can be used independently; however the greatest effect gives their joint application.

Введение

Перед руководством нефтяных компаний ежегодно возникают вопросы оптимального распределения инвестиций между месторождениями для обеспечения запланированных объемов нефтедобычи. Решение этих вопросов связано с трудоемкими процедурами расчета и согласования объемов добычи и необходимыми для этого финансовыми ресурсами. При этом возможны два основных варианта процесса согласования.

- 1) Высший менеджмент нефтяной компании формирует определенные значения объемов нефтедобычи дочерних акционерных обществ - ДАО, а те, в свою очередь, определяют необходимые суммы инвестиций, при которых эти объемы могут быть достигнуты. Если с точки зрения бюджета компании эти инвестиции приемлемы, то план по объему нефтедобычи и его финансирование оказывается согласованным. Если нет, то для ДАО предлагается скорректированный план объемов нефтедобычи и т.д., до тех пор, пока не будет получен согласованный с бюджетом всей компании план по нефтедобыче и его финансированию для всех ДАО.
- 2) ДАО предлагается подготовить предложения по вариантам объемов нефтедобычи и необходимых для реализации этих вариантов финансовых ресурсов. На основании таких прогнозов высшее руководство принимает решение о выборе программы финансирования нефтедобычи в целом по компании.

И в первом, и во втором случае приходится решать вопрос оптимизации финансирования программ нефтедобычи при ограниченных ресурсах.

В настоящее время выбор месторождений для выполнения плана по нефтедобыче и соответствующих инвестиций осуществляется, как правило, путем экспертных оценок. В лучшем случае проводится ранжирование месторождений по эффективности добычи и в соответствии с ним осуществляется распределение финансирования. В статье предлагается альтернативный подход на основе математического программирования и имитационного моделирования.

1 Исходные данные и критерии оптимизации

При принятии решений по финансированию добычи нефти и газа могут быть использованы разные критерии. В частности:

- при заданном объеме финансирования обеспечить **максимальный объем нефтедобычи** за планируемый период (полгода, год, два, три и т.д.).
- **для заданного объема** нефтедобычи **минимизировать** соответствующие **затраты** для планируемого периода.

Расчет и оптимизацию объемов финансирования и объемов добычи целесообразно проводить в разрезе месторождений, поскольку увеличивается число вариантов возможного распределения инвестиций, комбинируя которые можно найти более эффективное решение для компании в целом. При этом обеспечивается финансирование наиболее эффективных проектов (месторождений) и прекращается скрытое субсидирование неэффективных за счет эффективных.

Для расчета и оптимизации объемов инвестиций и объемов добычи **желательно (но не обязательно)** для каждого месторождения иметь несколько вариантов инвестиционных проектов.

2 Общий подход к оптимизации на основе статических и динамических моделей

Формирование вариантов оптимальных планов выполнения заданной программы добычи нефти связано с решением ряда задач. Это и задача оптимизации затрат для получения заданных объемов добычи, и задача оптимизации финансовых и материальных потоков во времени в рамках оперативного планирования. Эти задачи весьма сложны и трудоемки. Для их решения предлагается подход, состоящий из двух взаимосвязанных этапов.

На первом этапе методом, основанном на идее построения функции Беллмана [1-3] ищется решение, оптимизирующее количественные характеристики инвестирования для достижения заданных объемов нефтедобычи на соответствующих интервалах времени. Эти количественные характеристики инвестирования являются исходными данными для второго этапа.

На втором этапе решается задача оптимизации материальных и финансовых потоков во времени внутри временных интервалов. Варианты инвестиций и соответствующие им объемы добычи нефти, полученные на первом этапе, проверяются с точки зрения возможности реализации на втором этапе, в рамках динамической имитационной модели [4, 5]. Методологической основой здесь служат средства *непрерывного моделирования*, реализующие концепцию *системной динамики* Форрестера [6, 7]. В модели должны рассчитываться варианты доходов, получаемых от продажи нефти и размер их части, направляемой на инвестирование добычи (обратные связи, учет лагов), необходимые объемы заимствований в разные периоды времени, объемы основных выплат. На основании результатов динамического моделирования должны проводиться корректировки в исходных данных для модели первого этапа, уточняться общие объемы инвестиций и заимствований, с учетом выплат процентов по кредитам и прогнозов изменения цен на нефть.

3 Методика оптимизации затрат для выполнения планов добычи нефти в рамках статической модели

Задача. *Обеспечить заданный объем добычи нефти на каждом временном интервале с минимальными затратами на этом интервале.*

3.1 Целевая функция

Для каждого месторождения $n \in \{1, 2, \dots, N\}$ введем в рассмотрение вариант финансирования $j_{nt} \in \{1_{nt}, 2_{nt}, \dots, J_{nt}\}$ на временном интервале $t \in \{1, 2, \dots, T\}$. Каждый вариант j_{nt} опреде-

ляет объем финансирования X_{jnt} и соответствующий объем добычи V_{jnt} на месторождении n за период времени t . Полагаем $J_{nt} \geq 1$ при конечности общего числа вариантов финансирования.

Пусть $h_{jnt} = 1$, если вариант j_{nt} принимается для финансирования месторождения n на временном интервале t , иначе $h_{jnt} = 0$.

Тогда задача минимизации общих расходов на временном интервале t имеет вид:

$$(1) \quad S_t = \sum_{n=1}^N h_{jnt} X_{jnt} \rightarrow \min,$$

где минимум ищется среди всех возможных вариантов финансирования по каждому проекту $j_{nt} \in \{1_{nt}, 2_{nt}, \dots, J_{nt}\}$.

Решение задачи (1) – совокупность всех индексов h_{jnt} , должно удовлетворять следующим ограничениям:

$$(2) \quad \sum_{n=1}^N h_{jnt} V_{jnt} \geq V_t,$$

где V_t – лимит общей добычи нефти по всем месторождениям в каждом интервале времени t ;

$$(3) \quad \forall n, t \sum_{j_{nt}} h_{jnt} = 1$$

(т.е. для каждого месторождения в плановом периоде используется только один вариант инвестирования).

3.2 Алгоритм решения задачи (1-3)

1) Для решения задачи (1–3) условно будем считать, что на временном интервале t объем финансирования S_t изменяется от $\left(\min_{j,n} X_{jnt}\right)$ до $\left(\max_j \sum_{n=1}^N h_{jnt} X_{jnt}\right)$ с достаточно малым приращением Δ .

2) На первом шаге алгоритма запишем все имеющиеся варианты инвестирования и добычи для каждого месторождения в виде дискретной функции

$$3) \quad F(X) = \begin{cases} V_{jnt}, & \text{если } X = X_{jnt}; \\ 0 & \text{в противоположном случае.} \end{cases}$$

4) Далее для всех X из диапазона $\left[\left(\min_{j,n} X_{jnt}\right), \left(\max_j \sum_{n=1}^N h_{jnt} X_{jnt}\right)\right]$ проводим рекуррентные вычисления похожие на вычисления функции Беллмана, но с учетом ограничения (3):

$$(4) \quad B(X) = \max_{\chi} \{F(\chi) + B(X - \chi)\},$$

где $\chi = 0, \dots, X$.

Можно показать, что выражение (4) дает абсолютный оптимум добычи нефти на ресурсе X , если при его вычислении выполнено условие (3), т.е. инвестирование в месторождение либо производится только по одному из допустимых вариантов, либо вообще не производится.

Для выполнения условия (3) и сохранения глобального оптимума при каждом вычислении по (4) производится проверка участия месторождения, соответствующего добыче $F(\chi)$ в суммарной добыче определяемой $B(X - \chi)$. Если месторождение уже участвует для вычисления $B(X - \chi)$, то проверяется, увеличивает ли замена варианта инвестирования в это месторождение значение функционала $B(X)$ или нет. Если да, то производим замену варианта. Если нет, то переходим к другому месторождению, для которого возможны инвестиции в размере χ .

Для каждого значения X запоминается набор месторождений и вариантов инвестирования, дающих оптимальное значение.

Если для каждого месторождения *определен только один вариант инвестиций*, то процедура автоматически упрощается, так как проверяется только «повторное» инвестирование в месторождение.

Таким образом, если компания планирует добыть определенный объем нефти, который реален по геологическим условиям, а в проектах инвестиций месторождений есть вариант, предусматривающий максимальную добычу нефти, то плановый объем добычи будет среди получаемых по (4) вариантов или близок к некоторым из них. Более того, имея картину оптимальных объемов добычи нефти при разных вариантах финансирования, руководство может просто выбирать подходящий объем и тут же иметь перед глазами соответствующую оптимальную инвестиционную программу с учетом всех месторождений для взятого временного интервала t .

5 Методика динамического моделирования движения основных финансовых потоков

Задача динамической модели (в отличие от статических моделей - оптимизационных, эконометрических), предоставить возможность наблюдать за развитием процессов во времени, учитывать обратные связи, возникающие при движении материально-финансовых потоков, оценить влияние случайных возмущений, наконец, весьма наглядно отразить для пользователя моделируемые процессы и их взаимосвязи. Разумеется, подобную модель можно использовать и для оптимизации целевых показателей путем варьирования значений управляемых параметров.

Динамическое моделирование подразумевает имитацию движения материально-финансовых потоков на компьютере специальными программными средствами¹. Потоки имитируют интенсивность накопления и расхода финансовых средств, материальных запасов, направление этого накопления или расхода и т.д. Потоки, как правило, «вытекают» в накопители или «вытекают» из накопителей. Это позволяет оценивать объемы «накоплений» или «остатков» ресурсов. Таким образом, моделируется кругооборот основного и оборотного капитала компании, а также взаимоотношения с учреждениями финансово-кредитной сферы.

На основе динамического имитационного моделирования можно просматривать варианты финансовой политики и автоматически или вручную выбирать такие из них, которые позволяют:

- формировать финансовые резервы в объемах, обеспечивающих устойчивость компании;
- заранее планировать, когда, сколько и где необходимо денежных средств;
- определять, можно ли изыскать необходимые средства в рамках компании или придется обращаться к другим источникам и т.п.

6 Схема работы со статической и динамической моделями

- 1) На первом шаге составляются варианты инвестиционных проектов для каждого месторождения. Каждому варианту соответствует объем вложенных средств и соответствующий ему объем добычи нефти за плановый период. Увеличение числа вариантов проектов инвестирования в каждое месторождение позволит находить более оптимальную программу инвестиций в рамках всей компании. Но даже при наличии только одного варианта инвестирования для каждого месторождения возможна оптимизация.
- 2) Данные инвестиционных проектов являются исходными для работы статической оптимизационной модели. В результате работы модели находятся варианты максимальной добы-

¹ Например, пакета *iThink* компании High Performance Systems (<http://www.hps-inc.com>).

чи за плановый период для любого объема суммарных инвестиций S , а также, соответствующие этому объему оптимальные распределения инвестиций по месторождениям и ДАО.

- 3) Лицо, принимающее решение, выбирает привлекательный вариант объема добычи нефти и соответствующий ему объем суммарных инвестиций S . Этот выбор автоматически определяет проекты инвестиций в каждое месторождение. Если привлекательный вариант отсутствует, то следует перейти к пункту 1 и добавить варианты инвестиционных проектов для месторождений.
- 4) На динамической модели проверяется, как выбранный вариант будет реализовываться на практике. Какие потребуются объемы заимствований и в какое время, не произойдет ли вынужденное изменение объемов инвестирования из-за колебаний цен на нефть, насколько отклонится общая стоимость проекта от плановой и т.д.
- 5) Если моделирование выявило нежизнеспособность выбранной программы инвестиций, то осуществляется переход к пункту 3 и выбирается следующий вариант инвестиций.
- 6) Если динамическая модель показывает устойчивую работу компании для выбранной программы инвестиций, то на ее базе и на основе результатов имитации формируется и утверждается план инвестирования и объемы добычи нефти.

Заключение

В статье предложена методика решения задачи планирования инвестиций на примере нефтедобывающей компании, которая пригодна и для ряда других добывающих отраслей.

В основе методики лежит синтез оптимизационного и имитационного моделирования. Оптимизационная модель носит универсальный характер и основана на идеях построения функции Беллмана. Для построения динамической модели целесообразно использовать известные принципы системной динамики и соответствующие средства компьютерной имитации; эта модель представлена концептуально и должна разрабатываться для каждого конкретного случая. Оба подхода – статический и динамический, могут использоваться независимо, однако наибольший эффект дает их совместное применение.

Список литературы

- [1] Офицеров В.П., Судзиловский Н.Б. Об одном типе задач линейного программирования и их решении // Известия АН СССР. Техническая кибернетика. 1981. № 6. С. 14-17.
- [2] Офицеров В.П., Смирнов С.В. Моделирование и оптимизация снабженческой деятельности для крупных компаний // Проблемы управления и моделирования в сложных системах: Труды V международной конф. (Самара, июнь 2003). - Самара: СНЦ РАН, 2003. С. 197-205.
- [3] Беллман Р., Дрейфус С. Прикладные задачи динамического программирования. - М.: Наука, 1965.
- [4] Фролов Ю.В., Офицеров В.П., Ануфриев С.В. Моделирование и прогнозирование потребности педагогических кадров в городе Москве // Международный конгресс конференций «Информационные технологии в образовании». XIII Международная конференция «Информационные технологии в образовании»: Сб. трудов участников конференции. Часть V. – М.: Просвещение, 2003. С. 254-255.
- [5] Кугаенко А.А. Основы теории и практики динамического моделирования социально-экономических объектов и прогнозирования их развития. - М.: Вузовская книга, 1998.
- [6] Технология системного моделирования / Е.Ф. Аврамчук, А.А. Вавилов, С.В. Емельянов и др. – М.: Машиностроение; Берлин: Техник, 1988.
- [7] Форрестер Д. Основы кибернетики предприятия (индустриальная динамика): Пер. с англ. – М.: Прогресс, 1971.

УПРАВЛЕНИЕ КАЧЕСТВОМ ХИМВОДОПОДВОДКИ НА ТЭЦ

М.Ю. Лившиц¹, В.В. Солодянников²

¹Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
entcom@samgtu.ru

тел: +7 (8462) 32-11-06, факс: +7 (8462) 32-11-06

²ОАО и Э «Самараэнерго»
443100 г. Самара, Россия
ул. Маяковского 15, комн. 203
тел. +7 (8462) 42-08-47

Ключевые слова: идентификация, адаптация, автоматическое управление, химводоподготовка, математическая модель.

Abstract

It is determined to be discussed an effective approach to the synthesis of an adaptive control system by the process of chemical water admission technologies on a heat power plant using the model-identifier in a contour for the reception of the working information about the directly not measured parameters of the quality of water.

Введение

Проблемы получения химически очищенной и обессоленной воды для теплоэнергетических установок являются весьма актуальными в обеспечении надёжности и экономичности производства энергии. Обычно водоподготовительные установки содержат устройства предварительной очистки по удалению из воды примесей, взвесей, коллоидно-растворенных веществ разной степени дисперсности; устройства для дегазации воды и ионообменные установки (фильтры) для деминерализации воды путем эквивалентного обмена соответствующих ионов солей жесткости на ионы фильтрующего вещества с последующей периодической регенерацией фильтра.

Традиционные способы контроля и управления водоподготовительными установками становятся малопригодными для интенсивных технологий химводоподготовки (ХВО). Особенно это становится очевидным для теплоисточников, работающих в системах централизованного теплоснабжения с открытым разбором горячей воды населением и имеющих в тепловой схеме парогенераторы высокого давления. Схемы с открытым разбором горячей воды обуславливают высокие требования к её качеству согласно санитарным правилам, а надёжная работа парогенераторов высокого давления с большим значением теплового напряжения на поверхностях нагрева допускает использование добавочной воды лишь высокой степени обессоливания [1,2].

При этом, с учетом роста требования к защите окружающей природной среды от антропогенного воздействия, резкого повышения роли экономического стимулирования показателей процессов очистки воды, включая экологические, задачи контроля и управления работой отдельных элементов (особенно ионообменных фильтров) и всего технологического процесса химической очистки воды в целом, становятся весьма актуальными и выдвигаются на первый план.

Предлагается достаточно эффективный подход к решению этих задач путем синтеза адаптивной системы управления качеством ХВО с моделью-идентификатором в контуре.

1 Технологическое обоснование проблемы

Рассмотрим для определенности водоподготовительную установку подпитки теплосетей. Задача получения воды заданного качества для подпитки теплосетей (ПТС) может быть решена с помощью создания соответствующей системы управления [2]. Объектом управления в такой системе является участок технологической цепочки (рис. 1) от насоса исходной воды (НСВ) до бака декарбонизированной воды (БДВ), включающий Н-катионитные (Ф1-ФН), буферные (БФ1, БФ2) фильтры и клапаны – главный (ГКЛ) и подмеса (ПКЛ).

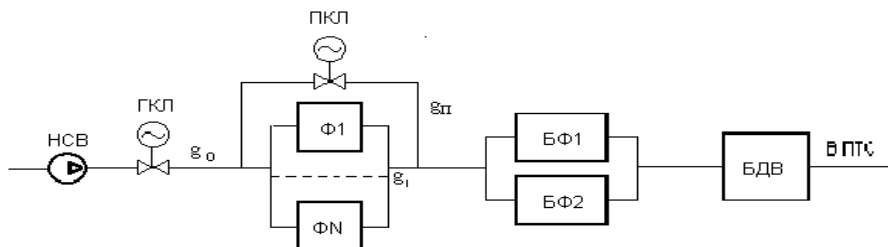


Рисунок 1 – Упрощенная технологическая схема объекта управления

Основным технологическим показателем качества воды является её карбонатный индекс

$$(1) \quad CI = Ca \cdot (HCO_3 + CO + OH),$$

где Ca , HCO_3^- , CO , OH^- - концентрации ионов Ca^{2+} , HCO_3^- , CO^{2-} , OH^- соответственно.

Карбонатный индекс CI зависит от соотношения расходов фильтрата g_1 и подмешиваемой к нему исходной воды $g_П$, так как именно этим соотношением определяются концентрации перечисленных ионов.

В типовых технологических ситуациях, например, в случае смешения различных сред в цилиндрическом смесителе может быть использована диффузионная модель в форме линейной краевой задачи:

$$(2) \quad \frac{\partial C_i(r, l, t)}{\partial t} = D_i \left(\frac{\partial^2 C_i(r, l, t)}{\partial l^2} + \frac{1}{r} \cdot \frac{\partial}{\partial r} \left(r \cdot \frac{\partial C_i(r, l, t)}{\partial r} \right) \right) - V_i(r) \cdot \frac{\partial C_i(r, l, t)}{\partial l} + F;$$

$$r \in (0, R_0), l \in (0, l_k),$$

$$t > 0; i = \overline{1, N} \quad \left. \frac{\partial C_i(r, l, t)}{\partial r} \right|_{r=R_0} = 0, \quad \left. \frac{\partial C_i(r, l, t)}{\partial r} \right|_{r=0} = 0, \quad \left. \frac{\partial C_i(r, l, t)}{\partial l} \right|_{l=l_k} = 0, \quad t > 0; i = \overline{1, N}$$

$$C_i(r, l, t)|_{t=0} = C_0, \quad C_i(r, l, t)|_{t=0} = C_0(r, l) \quad r \in [0, R_0], l \in [0, l_k].$$

Решение этой задачи, позволяет получить соответствующую объекту управления передаточную функцию[3]:

$$(3) \quad W_i(r, l, p) = \frac{C_i(r, l, p)}{C_{\text{вх}}(p)} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{8D_i}{V_i^2} \cdot \frac{K_{mi}(\gamma_m, l) \cdot K_{ni}(\beta_n, r)}{p \cdot (p+1)}.$$

Здесь K_{mi} , K_{ni} ядра соответствующих интегральных преобразований, $C_i(r, l, t)$ - концентрация i -го элемента; V_i - скорость конвективного массопереноса, D_i - коэффициент молекулярной диффузии.

Однако, система управления, основанная на принципе обратной связи по отклонению [4,5], не может быть реализована классическими методами. Это связано с тем, что карбонатный

индекс CI не измеряется непосредственно из-за отсутствия соответствующих измерительных приборов. Карбонатный индекс CI может быть вычислен, согласно (1), по результатам химического анализа потока. Экспериментальное определение многокомпонентных химических составов технологических потоков процессов химической водоочистки требует проведения специальных химических анализов в условиях реальных производственных процессов. Полное тщательное проведение химических анализов весьма трудоемко и не всегда экономически оправдано. Периодический и достаточно продолжительный процесс контроля основных качественных параметров ХВО путём химических лабораторных экспресс-анализов при невозможности прямого измерения этих параметров предполагает большое запаздывание в получении информации, высокую вероятность субъективных и объективных ошибок и, как следствие, неоперативное вмешательство в технологический процесс, снижение качества и надёжности ХВО. Таким образом, невозможность измерения карбонатного индекса в качестве управляемой величины определяет построение системы, в которой управляющее воздействие (расход $g_{II}(t)$) подмешиваемой к фильтрату воды), формируется в зависимости от отклонения рассчитанного по математическим моделям значения карбонатного индекса от его заданного значения. В этом случае структурная схема системы примет вид, представленный на рисунке 2, где $\bar{C}_0(t), g_0(t)$ - состав и расход исходной воды на входе; $g_{II}(t)$ - расход исходной воды на подмес - управляющее воздействие; CI^* - рассчитанное в блоке математического моделирования значение карбонатного индекса; $\varepsilon_{CI} = CI_{зад} - CI^*$ - сигнал рассогласования между рассчитанным в блоке математического моделирования значением карбонатного индекса CI^* и заданным $CI_{зад}$

Вычисление состава $\bar{C}_{XO}^*(t)$ химически очищенной воды, а значит и его карбонатного индекса CI^* , предполагает знание конструктивно-технологических параметров процесса химической очистки воды: геометрических параметров фильтра, характеристик фильтрующего материала, расходов основных технологических потоков процесса. В рассматриваемом случае процесс химической очистки воды подвергается влиянию многочисленных факторов, величину и характер которых весьма сложно учесть при моделировании. В частности, материал загрузки фильтров с течением времени теряет свои свойства, и после определенного числа фильтроциклов требует замены. Кроме того, процесс фильтрации на недавно загруженном фильтре отличается от такого же после нескольких месяцев работы. Эти обстоятельства обосновывают постановку адаптивной задачи управления объектом в условиях недостаточной априорной информации. При этом для восполнения недостающей априорной информации и построения работоспособной системы используются современные подходы теории

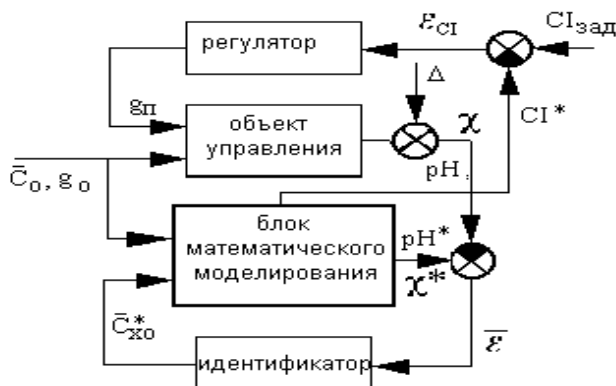


Рисунок 2 – Структурная схема системы управления качеством воды.

управления, связанные с активным использованием рабочей информации об объекте управления, т.е. информации, получаемой в ходе непосредственного функционирования объекта в рамках технологического процесса.

2 Постановка задачи идентификации

Оценку химического состава потока целесообразно производить по величинам, непосредственно зависящим от компонент состава, поддающихся прямому измерению, в качестве которых целесообразно использовать водородный показатель pH и электропроводность χ потока.

Таким образом, сформулирована задача определения неизмеряемого состояния объекта управления (его химического состава $\overline{C_{XO}}(t)$) по наблюдаемым величинам (pH и χ потока), т.е. задача идентификации состояния многомерного объекта управления [2,3].

Динамическую систему, которая формирует на выходе вектор состояния объекта $\overline{C^*}(t)$ по данным о его входах и выходах, будем в дальнейшем называть идентификатором состояния, если выполняется условие [4,5,6]:

$$(4) \quad \left| \overline{C_{XO}^*}(t) - \overline{C_{XO}}(t) \right| \rightarrow 0, \text{ при } t \rightarrow t_u < \infty.$$

Модель-идентификатор состоит из блока математического моделирования и идентификатора (см. рис.2). Для рассматриваемого случая под входом идентификатора будем понимать вектор $\overline{\varepsilon}$ отклонений рассчитанных значений pH^* и χ^* от доступных измерению величин pH и χ потока химически очищенной воды, а под выходом - определяемый вектор химического состава потока $\overline{C_{XO}^*}(t)$. На вход блока математического моделирования и объекта поступают в соответствующих масштабах одни и те же сигналы $\overline{C_0}(t)$ и $\overline{g_0}(t)$, следовательно, можно сравнить рассчитанные значения pH^* и χ^* соответствующие доступные измерению величины pH и χ , искаженные некоторой помехой Δ . Далее по результату сравнения идентификатор оценивает химический состав потока химически очищенной воды $\overline{C_{XO}^*}(t)$ для максимального приближения к действительному составу $\overline{C_{XO}}(t)$, который непосредственному измерению недоступен. После чего, по уточненному $\overline{C_{XO}^*}(t)$, рассчитывается по соотношению (1) карбонатный индекс CI^* , значение которого передается как действительное в контур управления расходом подмеса. При этом подлежит определению оптимальный способ оценки идентификатором состояния $\overline{C_{XO}^*}(t)$ объекта управления.

На практике, дуальный характер управления приводит к тому, что низкая скорость слабой сходимости в смысле некоторого функционала $J(\varepsilon)$ невязки между измеряемыми pH , χ и расчетными pH^* и χ^* значениями контролируемых величин за ограниченное время приведет не только к запаздыванию управляющего воздействия на время идентификации состояния, но и к погрешности оценки компонент вектора состояния химического состава потока. Поэтому получение заданного качества воды на выходе объекта возможно лишь при своевременном изменении расхода воды на подмес, т.е. при соблюдении принципа квазистационарности объекта [6]. В этих условиях по уточненному вектору состояния потока $\overline{C_{XO}^*}(t)$ удастся своевременно определить карбонатный индекс CI^* , который сравнивается с заданным значением

$CI_{зад}$, и сигнал ошибки ε_{CI} обрабатывается регулятором, изменяющим расход подмеса так, чтобы свести данный сигнал ошибки ε_{CI} к нулю.

Задача идентификации заключается в определении вектора состояния $\overline{C_{XO}^*}(t)$ с целью минимизации сигналов рассогласований $\overline{\varepsilon} = (\varepsilon_1; \varepsilon_2)$.

3 Формальный алгоритм адаптивной системы управления

Зададимся целью формального построения модели-идентификатора, обеспечивающей определение качественных и количественных характеристик технологического процесса по косвенным измерениям, а также подсистемы адаптации этой модели, включающей прогнозирование результатов действий оператора и технологических операций и диагностику состояния системы управления. Для этого введем следующие понятия.¹

1) Пусть вектор состояния системы $X = (x_1, x_2, \dots, x_N)$ определен над пространством $R^N : N < \infty$, которое будем называть пространством состояний системы. В условиях технологических, энергетических и материальных ограничений состояние системы X может быть допустимым или недопустимым. Область $D \subseteq R^N$ будем называть допустимой областью состояний системы, если она содержит все допустимые состояния системы, и не содержит никаких других.

Множество функций $H = (h_i : R^N \rightarrow R^1 | i = \overline{1, n})$ будем называть множеством функций ограничения состояний системы, а множество функций $F = (f_j : R^N \rightarrow R^1 | j = \overline{1, m})$ – множеством функций структурной связи состояний системы, если

$$\forall X \in D \Leftrightarrow (\forall_i : 1 \leq i \leq n \Rightarrow h_i(X) > 0) \wedge (\forall_j : 1 \leq j \leq m \Rightarrow f_j(X) = 0)$$

Таким образом, множества функций ограничения H и функций структурной связи F задают допустимую область D , и тем самым отражают априорные сведения о рассматриваемой системе и происходящих в ней процессах.

Функцию $w : R^N \cdot R^K \cdot R^1 \rightarrow R^N$ будем называть функцией идентификации состояния системы по функциональному критерию качества $J \in \Omega = (R^N \cdot R^K \rightarrow R^1) : K < N$, если $w(X, \varepsilon, h) \equiv \arg \min_{X^* \in D} J(X^*, \varepsilon)$, где $X \in D$ – опорное состояние системы (начальная точка

поиска); ε – вектор невязки косвенных измерений; h – радиус окрестности поиска. Моделью-идентификатором M на этом этапе будем называть четверку математических объектов $M = (H, F, J, X_0)$, где $X_0 \in D$ – начальное состояние системы. Тогда, модельная траектория $X(t)$ движения системы с дискретным временем t есть последовательность состояний системы $X(t) = (X_t = w(X_{t-1}, \varepsilon_{t-1}, h))_{t=1,2,3,\dots}$, а адаптация модели есть конкатенация траектории с новым начальным состоянием системы.

Начальное состояние системы должно удовлетворять множествам функций ограничения H и функций структурной связи F . В силу неточности методов измерений, субъективных и объективных ошибок при определении результатов измерений, обычно такое состояние системы может оказаться недопустимым, т.е. $X_0 \notin D$, и поэтому нуждается в целесообразной корректировке. Для этого процесс задания начального состояния системы разделим на два этапа:

¹ Результаты получены совместно со студентами А.Афанасьевым и В. Пироцкой.

а) процесс последовательного поступления информации; б) комплексный анализ поступившей информации в целом. Формализуем это следующим образом.

Рассмотрим информационный канал с дискретным источником I генерирующим поток данных в форме сообщений о присвоении параметру состояния x_i значения a . В этом случае сообщение образует упорядоченную пару чисел $(i, a) \in \overline{(1, N)} \cdot \overline{R^1}$, обозначать которую будем синтаксической конструкцией вида $x_i = a$.

Тогда, поток данных Q_I , генерируемый источником I , есть последовательность сообщений $Q_I = (x_{i_k} := a_k)_{k=1,2,3,\dots}$.

Вообще говоря, каждое сообщение содержит еще некоторую информацию, связанную со знанием свойств технологического процесса, в той или иной степени уже использованных при построении множеств функций ограничения H и функций структурной связи F . Эту информацию можно получить (синтезировать) путем построения соответствующего анализатора (прогноза), генерирующего дополнительные сообщения.

Функцию $g_{x_i}^{x_i} : R^N \cdot R^1 \rightarrow R^1$ будем называть функцией прогноза параметра состояния системы $x_j : 1 \leq j \leq N$ по параметру состояния системы $x_i : 1 \leq i \leq N$, если $x_i := a_k \Rightarrow x_j := g_{x_i}^{x_j}(\tilde{X}_k, a_k)$, где \tilde{X}_k – состояние системы в момент получения k -го сообщения.

Множество функций прогноза $G = (g_{x_i}^{x_j} | (i, j) \in \overline{(1, N)} \cdot \overline{(1, N)})$ будем называть прогнозом потока данных Q_I .

Прогноз потока данных необходимо реализовать на первом этапе задания начального состояния системы. Второй этап будет заключаться в интегральной поправке.

Функцию $r_{D^*} : D^* \rightarrow R^N$ будем называть функцией поправки по области $D^* : R^N \supseteq D^* \supset D$, если $(\forall X \in D^* / D \Rightarrow r_{D^*}(X) \in D) \wedge (\forall X \in D \Rightarrow r_{D^*}(X) = X)$. Если $D^* = R^N$, $D^* = R^N$, то функция r_{D^*} является функцией полной поправки (см. рис. 3).

Таким образом, алгоритм S задания начального состояния системы есть последовательная выборка сообщений из потока данных Q_I с прогнозом G и последующей поправкой r_{D^*} . Далее, построим механизм диагностики системы управления, характеризующий положение системы в своем пространстве состояний.

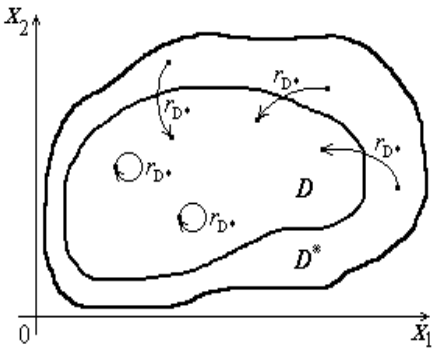


Рисунок 3 – Функция поправки.

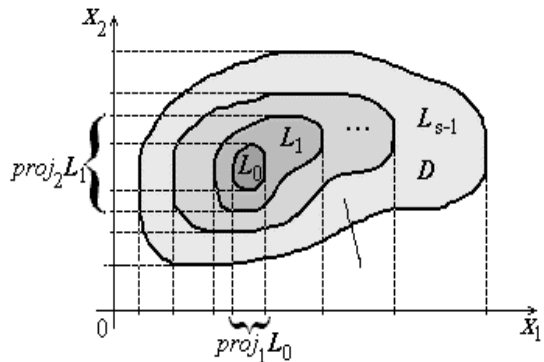


Рисунок 4 – Множество критических уровней.

Множество областей

$$(5) L = \left\{ L_k \mid k = \overline{0, s-1}, (\forall i \neq j : 0 \leq i, j < s \Rightarrow L_i \cap L_j = \emptyset) \wedge \left(\bigcap_{k=0}^{s-1} L_k = D \right) \right\}$$

будем называть множеством критических уровней. Функцию $T : R^N \rightarrow \prod_{k=1}^N \{0, s\}$ будем называть

функцией теста состояния системы по множеству критических уровней L , если $\left[(\forall X \in D \Rightarrow T(X) = (a_1, a_2, \dots, a_N)) \Leftrightarrow (\forall i : 1 \leq i \leq N \Rightarrow a_i = \min \{ a^* : x_i \in \text{proj}_i L_{a^*} \}) \right] \wedge \left[\forall X \notin D \Leftrightarrow T(X) = (s, s, \dots, s) \right]$.

Определенная таким образом функция теста T дает покомпонентное распределение состояния системы X по критическим уровням L . Заметим, что теперь определение модели - идентификатора M можно расширить до восьмерки математических объектов $M = (H, F, J, X_0, Q_H, G, r_{D^*}, L)$.

4 Конструктивная проблемно-ориентированная математическая модель ХВО¹

Для системной структуризации математической модели ХВО охарактеризуем состояние системы тремя потоками: исходной воды, $g_0(t)$, фильтрата $g_1(t)$, химочищенной воды $g_n(t)$. Компонентами вектора состояния каждого потока $C = (G, p, t, Ca, Mg, Na, H, Fe, SO_4, Cl, OH, HCO_3, CO, SiO, CO_2)$ являются: расход G потока (т/час), давление p в потоке (Па), температура t потока ($^{\circ}C$), концентрация Ca катиона кальция Ca^{2+} (мг-экв/л), концентрация Mg катиона магния Mg^{2+} (мг-экв/л), концентрация Na катиона натрия Na^+ (мг-экв/л), концентрация H катиона водорода H^+ (мг-экв/л), концентрация Fe катиона железа Fe^{2+} (мг-экв/л), концентрация SO_4 сульфат-иона SO_4^{2-} (мг-экв/л), концентрация Cl хлор-иона Cl^- (мг-экв/л), концентрация OH гидроксил-иона OH^- (мг-экв/л), концентрация HCO_3 бикарбонат-иона HCO_3^- (мг-экв/л), концентрация CO карбонат-иона CO_2^- (мг-экв/л), концентрация SiO аниона кремниевой кислоты SiO_2^- (мг-экв/л), концентрация CO_2 углекислоты CO_2^{3-} (мг-экв/л).

Таким образом, получили восьмерку математических объектов $(H, F, J, X_0, Q_H, G, r_{D^*}, L)$, образующих модель-идентификатор M , а вектор состояния системы X примет вид: $X = (C_u | C_\phi | C_{xo}) \in R^{45}$, где C_u – исходная вода; C_ϕ – фильтрат; C_{xo} – химочищенная вода; R^{45} – сорока пяти - мерное евклидово пространство состояний системы. В составе модели:

- 1) Множество функций ограничения H состояний системы выделяют прямоугольный гиперпараллелепипед в пространстве состояний, представляющий собой возможные диапазоны изменения физических параметров потоков, концентраций химических элементов и их соединений.
- 2) Множество функций структурной связи F включает в себя [1,2,3]:
 - а) f_1 – химический потенциал потоков. Связь выделяет множество химически нейтральных потоков, т.е. определяет возможность химического взаимодействия в потоке, являющегося результатом физического смешения двух других потоков.

¹ Результаты получены совместно с доцентом Ю.Э. Плешивцевой.

- b) f_2 – электрический потенциал потоков. Функция определяется разностью катионов и анионов в каждом потоке: сумма катионов равна сумме концентраций, Ca, Mg, Na, H, Fe , а анионов – сумме концентраций SO, Cl, OH, HCO, CO, SiO . Связь выделяет множество электронейтральных потоков (неэлектронейтральные потоки не могут существовать в стационарном состоянии).
- c) f_3, f_4, f_5, f_6 – Соотношения концентраций ионов кальция Ca и магния Mg между собой в потоках.
- d) f_7 . Связь предполагает, что концентрации некоторых ионов в потоке малы и их можно считать равными нулю, например: $OH = CO = 0$.
- e) f_8 . Связь отражает зависимость типа реакции раствора (кислотная или щелочная) от концентрации ионов водорода H или гидрокарбона-иона HCO_3 (исходная вода дает щелочную реакцию).
- f) $f_9 : Na_u = Na_\phi = Na_{xo}; f_{10} : SO_{4u} = SO_{4\phi} = SO_{4xo}; f_{11} : Cl_u = Cl_\phi = Cl_{xo};$
 $f_{12} : SO_{4u} = SO_{4\phi} = SO_{4xo}; f_{13} : Fe_u = Fe_\phi = Fe_{xo}.$
- g) Данные связи отражают тот факт, что в процессе фильтрации концентрации Na, SO_4, Cl, SiO, Fe не изменяются, что так же является особенностью процесса ионного обмена в \square -катионитном фильтре, связанные с типом загруженного катионита.

3) В качестве критерия качества идентификации J выберем квадратичный функционал:

$$J = \left(\frac{pH_u - pH_u^*}{pH_u} \right)^2 + \left(\frac{\chi_u - \chi_u^*}{\chi_u} \right)^2 + \left(\frac{pH_\phi - pH_\phi^*}{pH_\phi} \right)^2 + \left(\frac{\chi_\phi - \chi_\phi^*}{\chi_\phi} \right)^2$$

где $pH_u, pH_u^*, pH_\phi, pH_\phi^*, \chi_u, \chi_u^*, \chi_\phi, \chi_\phi^*$ – расчетные и измеренные значения водородного показателя и электропроводности исходной воды и фильтрата соответственно.

Функция идентификации w состояния системы реализована путем численной оптимизации критерия J , характеризующего невязку ε . В качестве метода поиска предлагается многомерная модификация метода Ньютона.

- 4) Начальное состояние системы X_0 определяется проведением химического анализа потоков с использованием алгоритма S задания начального состояния системы.
- 5) Поток данных Q_{II} образуется вводом химического анализа потоков.
- 6) Прогноз G потока данных Q_{II} состоит из функций прогноза параметров состояния системы, полученных по результатам работы [2], причем функция поправки r_{D^*} образована композицией двух других функций $r_{D^*} = r_2(r_1(X))$, где функция r_1 компенсирует электронейтральность каждого потока за счет концентрации хлора CL , а функция r_2 приводит потоки в состояние химического равновесия [1,2,3].
- 7) Множество критических уровней L состоит из двух областей, т.е. $s = l : s = l$: уровень L_0 , ограниченный прямоугольным гиперпараллелепипедом допустимости, представляющим собой предусмотренные технологическими нормами диапазоны изменения параметров потоков, концентраций химических элементов и их соединений; и уровень L_1 , образованный разностью допустимой области D состояний системы и уровнем

$$L_0, \text{ т.е. } L_1 = \frac{D}{L_0}.$$

Таким образом, функция T теста состояния системы показывает принадлежность параметров состояния потока допустимым технологическими нормами, физически возможным, и физически невозможным диапазонам.

5 Экспериментальная проверка результатов.

Предложенные здесь алгоритмы реализованы и доказали свою высокую эффективность в ходе промышленной эксплуатации в течение ряда лет на крупной ТЭЦ. Для оценки степени соответствия модели реальной системе были проведены исследования с использованием данных химических анализов воды цеха химводоподготовки Самарской ТЭЦ. Для экспериментальной оценки качества работы наблюдателя, были выделены базовые составы фильтрата, соответствующие типовым значениям концентраций химических элементов в нём, а вся совокупность данных была разбита на выборки, в каждой из которых концентрации элементов, от которых существенно зависит значение pH и χ , значительно отличаются от базовых значений.

Таким образом, были получены экспериментальные выборки соответственно. Для каждой из них определены расчётные значения pH^* и χ по математической модели и относительное отклонение ΔpH_i^* и χ .

В таблицу 1 сведены результаты статистической обработки: математическое ожидание M отклонения pH и χ ; максимальное относительное отклонение $\max \Delta pH$ и $\max \Delta \chi$; дисперсии отклонения $D(\Delta pH)$ и $D(\Delta \chi)$; среднеквадратичное отклонение, $\nabla(\Delta pH)$ и $\nabla(\Delta \chi)$; объем выборки N .

На рисунках 5 и 6 приведены результаты работы идентификатора по критериям сравнения расчетных значений водородного показателя (рис 5) и электропроводности (рис.6) с экспериментальными для представительных выборок по химсоставу потоков.

Таблица 1 – Статистические характеристики

	№ выборки	M	$\max \Delta$	D	$\nabla(\Delta)$	N
Водородный показатель pH	1	4,51	10,56	8,12	2,85	76
	2	1,01	8,17	12,37	3,52	58
	3	0,98	9,84	17,57	4,19	31
	4	17,11	22,64	8,77	2,96	21
Электропроводность χ	1	-15,05	-4,73	27,27	5,22	21
	2	-9,56	7,64	51,38	7,17	58
	3	-0,54	9,86	36,57	6,05	18
	4	-1,02	12,87	80,49	8,97	30
	5	-8,45	7,64	68,12	8,25	41
	6	-16,35	-5,53	31,76	5,64	19
	7	-2,81	7,64	32,44	5,70	22
	8	-7,48	4,48	36,98	6,08	33
	9	0,36	12,87	69,53	8,34	27
	10	-3,76	9,88	43,39	6,59	34
	11	0,93	16,72	49,93	7,07	21
	12	1,64	9,88	47,29	6,88	37

Экспериментальная проверка математических моделей на статистически представительных выборках показала, что разница между значениями параметров, получаемых расчетом и экспериментально, не превышает 10%.

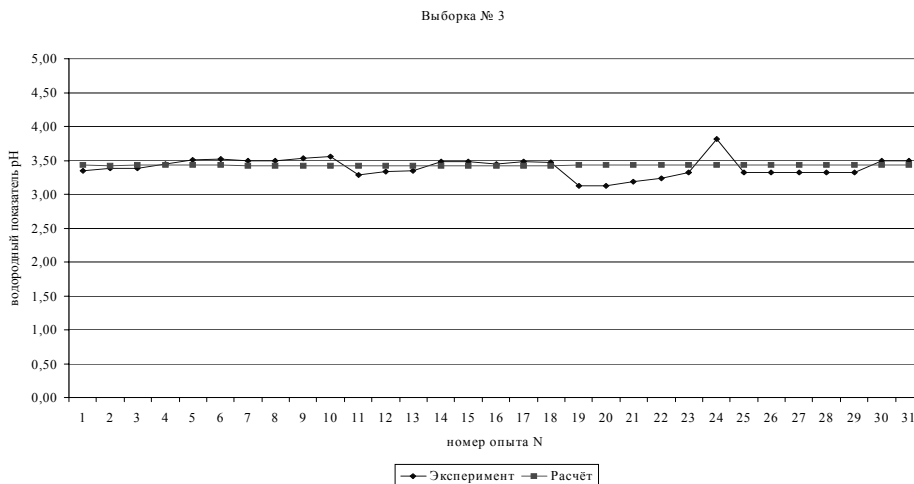


Рисунок 5 – Экспериментальная проверка идентификатора по водородному показателю.

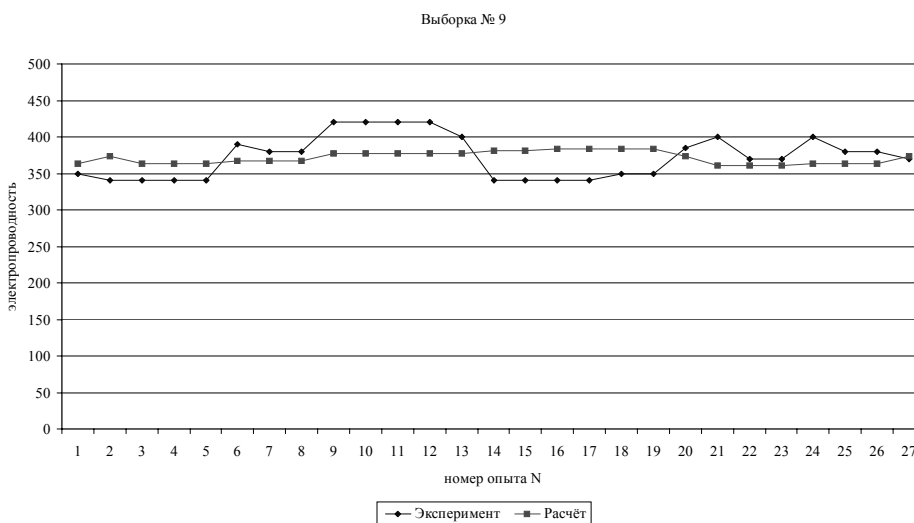


Рисунок 6 – Экспериментальная проверка идентификатора по электропроводности

Заключение

Опыт эксплуатации системы позволяет констатировать удовлетворительное качество идентификации химсостава и управления по качественным характеристикам ХВО. На рисунке 7 приведен для примера рабочий фрагмент мнемосхемы АСУТП.

ХИМСОСТАВ										
	ИСХОДНАЯ ВОДА		НкФ-1		НкФ-2		НкФ-3		ХИМ. ОЧИЩ. ВОДА	
	ИЗМЕР.	РАССЧ.	ИЗМЕР.	РАССЧ.	ИЗМЕР.	РАССЧ.	ИЗМЕР.	РАССЧ.	ИЗМЕР.	РАССЧ.
Время химанализа	12:05:44 15-11-99		12:05:54 15-11-99		12:06:15 15-11-99		12:06:21 15-11-99		12:06:39 15-11-99	
Расход	459		0		138		149		468	
pH		5.89	3.49	0.00	3.46	3.40	3.68	3.41	5.52	4.31
Э		383	382	0	367	363	372	360	329	287
Ж	3.600	3.600	1.600	0.000	1.300	1.072	1.300	1.082	2.460	2.099
H (Кисл.)	0.000	0.000	0.000	0.000	0.200	0.428	0.200	0.418	0.000	0.000
HCO (Щ)	2.100	2.100	0.100	0.000	0.000	0.000	0.000	0.000	0.850	0.599
Na	15.0	15.0	15.0	0.0	15.0	22.5	15.0	22.5	15.0	19.5
Ca	2.400	2.400	1.067	0.000	0.867	0.715	0.867	0.721	1.770	1.399
Mg	1.200	1.200	0.533	0.000	0.433	0.357	0.433	0.361	0.690	0.700
SO ₄	86.4	84.6	86.4	0.0	86.4	100.3	86.4	100.3	86.4	93.9
Cl	24.0	9.5	9.5	0.0	9.5	9.5	9.5	9.5	9.5	9.5
SiO	5.0	5.0	5.0	0.0	5.0	5.0	5.0	5.0	5.0	5.0
Fe	280.0	280.0	280.0	0.0	280.0	280.0	280.0	280.0	280.0	280.0
CO ₂	5.6	5.6	93.9	0.0	98.3	98.0	98.3	98.0	60.0	61.1
ОН	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Карбонатный индекс		Предыдущая		ЛАБОРАНТ:	
Заданный	1.00	подстройка модели		Тюрина	
Расчетный	0.84	12:06:54 15-11-99		ВСТ	
12:06:51 Введен хим. анализ хим. очищенной воды				12:08:07	
12:06:54 Произведен пересчет карб. индекса					

Рисунок 7 – Фрагмент мнемосхемы АСУТП ХВО

Список литературы

- [1] Солодяников В.В., Чупрунов В.П., Лившиц М.Ю., Муранов С. В., Щербаков Е.В. АСУ подготовкой воды на ТЭЦ //Промышленные АСУ и контроллеры №5 2000, с. 1-6
- [2] Солодяников В.В. Расчет и математическое моделирование процессов водоподготовки. М.: Энергоатомиздат, 2003г.
- [3] Лившиц М.Ю., Солодяников В.В. АСУ ТП водоподготовки. //Труды международной конференции "Надежность и качество в промышленности, энергетике и на транспорте". Самара, 1999, с. 174-175.
- [4] Цыпкин Я.З. Адаптация и обучение в автоматических системах М, Наука, 1968.
- [5] Эйхофф, П. Основы идентификации систем управления, М, Мир, 1975
- [6] В.А. Терехов, И.Ю. Тюкин Адаптивные системы управления: проблемы и тенденции //Управление и информационные технологии. Всероссийская научная конференция. Сборник докладов. Т.1., Санкт-Петербург, 2003г. с.146-154.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ УПРАВЛЕНИЯ ВЕНТИЛЯЦИОННЫМ РЕЖИМОМ СЕВЕРО-МУЙСКОГО ТОННЕЛЯ

В.В. Тарасевич¹, А.А. Мороз², В.В. Окольнішников³

¹Новосибирский государственный архитектурно-строительный университет «Сибстрин»
630008, Новосибирск, 8, ул. Ленинградская, 113, Россия
tvv@iis.nsk.su

тел. +7 (3832) 66-47-87, факс: +7 (3832) 66-40-83

²ФГУП «СибНИА им. С.А. Чаплыгина»
630051, Новосибирск, ул. Ползунова, 21, Россия
moroza@mail.cis.ru

тел. +7 (3832) 59-59-17; факс +7 (3832) 79-54-25

³КТИ вычислительной техники СО РАН
630090, Новосибирск, 90, ул. Институтская, 6, Россия
okoln@kti.nsc.ru

тел.: +7 (3832) 30-25-72, факс: +7 (3832) 34-43-61

Ключевые слова: тоннель, вентиляция, тепловой режим, управление, расчет.

Abstract

The mathematical model of ventilation of Northern Muya Railway Tunnel in various climatic conditions is considered; the winter season is most hard here. The control system of the ventilation mode is intended for the maintenance of positive temperatures in the tunnel in winter time and avoidance of an icing. The mathematical model describing a non-stationary ventilating mode, arising in the tunnel at a train entrance and moving along tunnel is submitted in the paper. Application of such model allows to reveal "bottlenecks" dangerous and to select the most effective modes. Results of work will provide the simulation of the influence of climatic and operational factors on ventilating mode Northern Muya Railway Tunnel and an estimation of the efficiency of various control facilities of its parameters

Введение

В больших железнодорожных тоннелях, особенно в таких гигантских сооружениях, как Северо-Муйский тоннель, расположенный в суровой климатической зоне, необходимо в зимнее время поддерживать положительные температуры на протяжении всей трассы тоннеля, во избежание его обледенения. Для этой цели тоннель оборудован группами вентиляторов и калориферов, обеспечивающих подачу в тоннель теплого воздуха. Управление включением, выключением и изменением подачи теплового воздуха путем подключения или выключения дополнительных калориферов и вентиляторов осуществляется в зависимости от показаний датчиков, расположенных по трассе тоннеля. Входной и выходной порталы тоннеля закрыты воротами, которые открываются перед прохождением поезда.

Оптимальное управление режимом вентиляции, т.е. поддержание требуемого микроклимата в тоннеле при минимальных затратах энергии, существенно снижает эксплуатационные издержки. Для нахождения наиболее эффективных режимов вентиляции необходимо опираться на достаточно полную и адекватную математическую модель вентиляции тоннеля.

Наибольшие возмущения теплового и аэродинамического режимов тоннеля возникают при входе и прохождении поезда. Здесь требуется своевременное включение и выключение вентиляторов в зависимости от динамических параметров потока, которое должно ориентироваться не только на апостериорные показания датчиков, но и работать на опережение. Этого

можно добиться на основе математической модели, позволяющей рассчитывать и предсказывать последствия того или иного управляющего воздействия.

Обычно для расчетов вентиляционных режимов в тоннелях применяется относительно простая модель несжимаемой жидкости [1], которая упрощенно учитывает влияние давления и температуры на плотность воздуха; при этом поле температур определяется на основе рассчитанных предварительно гидравлических параметров потока как решения вспомогательной задачи.

В данной работе предлагается более точная модель, учитывающую термогидравлические параметры потока (давление, скорость, температура), зависимость плотности, вязкости и т.п. параметров воздуха от давления и температуры, теплообмен через стенки тоннеля и др. Поле давления, скоростей и температур насчитывается одновременно как взаимосвязанные параметры потока [2].

В дальнейшем будет применяться следующая терминология: железнодорожный тоннель, сбойки, порталы, второй ствол и другие протяженные участки тоннеля будут именоваться каналами тоннеля; места расположения различной арматуры (вентиляторы, калориферы и т.п.), местные сопротивления (ворота), соединения каналов (ответвления, пересечения и т.п.) будут именоваться узлами тоннеля.

1 Область определения

Система узлов и каналов тоннеля, по которым происходит перемещение воздуха, будет описываться одномерным комплексом, или ориентированным графом с приписанной каждому ребру длиной [2]. Ориентация ребер будет означать положительное направление потока. Такой граф будет выступать в качестве области определения задачи, и представляет собой обобщение оси x .

Пусть каналы и узлы тоннеля пронумерованы, при этом i – номер канала, j – номер узла. Величины, относящиеся к каналам, будут снабжаться нижним индексом i , а величины, относящиеся к узлам – верхним индексом j .

Функция инциденций η_i^j определяется равной: $+1$, если i -ый канал входит в узел j ; -1 , если i -ый канал выходит из узла j ; 0 – в остальных случаях.

В качестве основных параметров потока воздуха принимаются $p = p(x, t)$ – давление; $T = T(x, t)$ – абсолютная температура, $V = V(x, t)$ – средняя по сечению скорость, где x – координата вдоль тоннеля; t – время.

2 Основные уравнения

2.1 Применяемые допущения

Так как рабочая область температур и давлений лежит весьма далеко от критической точки, то, без особых погрешностей, воздух можно считать идеальным газом [3], имеющим уравнение состояния:

$$(1) \quad \rho = \frac{P}{RT},$$

где ρ – плотность воздуха, R – газовая постоянная (для воздуха $R = 287$ дж/кг·К [3]).

Согласно экспериментальным данным [3], в диапазоне температур от -50°C до $+50^\circ\text{C}$ удельная теплоемкость при постоянном давлении c_p для воздуха изменяется только на 0,33%, а показатель адиабаты $\gamma = c_p/c_v$ – на 0,07%; поэтому эти физические характеристики будем считать постоянными, полагая: $c_p \approx 1003$ дж/кг·К, $\gamma \approx 1,4$.

2.2 Уравнения движения воздуха в тоннеле

Нестационарное неизотермическое движение воздуха в каждом i -том канале тоннеля (индекс i далее для простоты опущен) в условиях теплообмена с окружающей средой, в случае существенно дозвукового потока (что будет иметь место в действительности), можно описывать обобщенными уравнениями гидравлического удара [2] для идеального газа:

$$(2) \quad \frac{\partial p}{\partial t} + V \frac{\partial p}{\partial x} + \rho a^2 \frac{\partial V}{\partial x} = \frac{\gamma R Q_w}{\omega c_p},$$

$$(3) \quad \frac{\partial V}{\partial t} + V \frac{\partial V}{\partial x} + \frac{1}{\rho} \frac{\partial p}{\partial x} = - \frac{f_w}{\rho \omega} - g \frac{dz}{dx},$$

$$(4) \quad \frac{\partial T}{\partial t} + V \frac{\partial T}{\partial x} + \frac{a^2}{c_p} \frac{\partial V}{\partial x} = \frac{\gamma Q_w}{\rho \omega c_p},$$

где ω – площадь поперечного сечения потока в тоннеле; g – ускорение силы тяжести; $z=z(x)$ – вертикальная отметка дна тоннеля.

Здесь $Q_w = q_w \cdot \chi$ – удельный тепловой приток извне на единицу длины тоннеля, или линейная плотность теплового потока [4]; $f_w = \tau \cdot \chi$ – удельная сила сопротивления трения на единицу длины тоннеля, где τ – касательное напряжение трения о стенки, χ – смоченный периметр, a – скорость распространения малых возмущений в жидкости (скорость звука):

$$(5) \quad a = \sqrt{\gamma RT}$$

2.3 Сопротивление трения

Удельное сопротивление трения на участке тоннеля без поезда может быть найдено из соотношений:

$$(6) \quad f_w = \tau \chi = \lambda \frac{\rho |V| V}{8} \chi,$$

где τ – касательное напряжение трения на стенке тоннеля; χ – периметр поперечного сечения тоннеля; λ – коэффициент гидравлического трения [5].

Удельное сопротивление трения на участке тоннеля с поездом может быть найдено из соотношений:

$$(7) \quad f_w = \tau \chi + \tau_n \chi_n = \lambda \frac{\rho |V| V}{8} \chi + \lambda_n \frac{\rho |V_n - V| \cdot (V_n - V)}{8} \chi_n,$$

где первое слагаемое определяется по предыдущим формулам; V_n – скорость движения поезда; τ_n – касательное напряжение трения на стенке поезда; χ_n – периметр поперечного сечения поезда; λ_n – коэффициент гидравлического трения воздуха о стенки поезда [6, 7].

Движение поезда задается графиком скорости $V_n = V_n(t)$.

2.4 Тепловые потоки

В общем случае, плотность теплового потока q_w определяется из решения внешней тепловой задачи.

Упрощенно, плотность теплового потока q_w можно в первом приближении определять по формуле [4, 8, 9]:

$$(8) \quad q_w = \frac{T_{env} - T}{R_T} = K \cdot (T_{env} - T),$$

где R_T – термическое сопротивление; $K=1/R_T$ – коэффициент теплопередачи; T_{env} – температура внешней среды.

Различные формулы для определения K приведены в работах [8, 9].

3 Граничные условия и начальные данные

Для замыкания задачи следует задать еще граничные условия и начальные значения параметров потока.

3.1 Граничные условия на неподвижных границах

Граничные условия на неподвижных границах представляют собой математические модели оборудования и узлов сопряжения (входной и выходной порталы, соединения штолен, вентиляционное оборудование и др.).

3.1.1 Портал

Входной и выходной порталы будут описываться соотношениями

$$(9) \quad p_{env} - p^j = S^j \left| \eta_i^j G_i^j \right| G_i^j,$$

где p_{env} – давление наружной среды (атмосферное); S^j – сопротивление входя/выхода портала; $G_i^j = \rho_i^j \omega_i V_i^j$ – массовый расход.

При условии втекания воздуха в портал (т.е. когда $\eta_i^j G_i^j < 0$) ставится дополнительное условие

$$(10) \quad T_i^j = T_{атм},$$

где $T_{атм}$ – температура наружного воздуха.

3.1.2 Разветвление каналов

При разветвлении каналов ставятся условия:

Баланс массовых расходов в узле описывается соотношением

$$(11) \quad \sum_{i \in R^j} \eta_i^j G_i^j = 0,$$

в узле задается равенство давлений (сопротивлениями в узле пренебрегаем):

$$(12) \quad p_i^j = p^j, \quad \forall i \in R^j$$

Из условия баланса тепловой энергии в узле, полагая, что там происходит полное перемешивание сливающихся потоков, получаем

$$(13) \quad \sum_{i \in R_n^j} \left[\bar{c}_{p,i} (T_i^j - T^j) \right] \eta_i^j G_i^j = 0,$$

где сумма берется только по вытекающим в узел потокам, т.е. когда $\eta_i^j G_i^j > 0$; T^j – температура в узле.

Температура всех вытекающих потоков (т.е. когда $\eta_i^j G_i^j \leq 0$) полагается равной температуре в узле:

$$(14) \quad T_i^j = T^j$$

Соотношение (13) позволяет определить температуру в узле T^j .

3.1.3 Вентиляторы и калориферы

Калориферы, объединенные с вентиляторами, описываются соотношениями:

$$(15) \quad \Delta p^j = f_G(G, t, u),$$

где Δp – перепад давления, создаваемый на вентиляторе, G – расход, перекачиваемый вентилятором; $f_G(G, t, u)$ – напорно-расходная характеристика вентилятора; u – управляющий сигнал.

$$(16) \Delta T^j = f_T(G, t, u),$$

где ΔT – перепад температуры на вентиляторе, $f_T(G, t, u)$ – тепловая характеристика вентилятора; u – управляющий сигнал, определяющий режим его работы.

3.2 Условия на подвижной границе (оконечности поезда)

На торцах движущегося поезда ставятся условия баланса массы и равенства температур:

$$(17) [\rho\omega(V - V_n)] = 0, \quad [T] = 0,$$

где знак $[f]$ означает разность значений величины f перед и за фронтом передвижения поезда, и задается закон сопротивления:

$$(18) [p] = -\rho\zeta_n |V - V_n| (V - V_n) / 2,$$

где ζ_n – сопротивление оконечности поезда.

3.3 Начальные значения

В качестве начальных значений параметров потока задается распределение давлений, скоростей и температур $p_{i,0}(x)$, $V_{i,0}(x)$, $T_{i,0}(x)$ в начальный момент времени ($t=0$) по всем каналам (стволом и штольням) тоннеля:

$$(19) p_i(x, 0) = p_{i,0}(x), \quad V_i(x, 0) = V_{i,0}(x), \quad T_i(x, 0) = T_{i,0}(x),$$

где i – номер канала, а также параметры всех тепловых потоков и значения параметров всех узлов, входящих в тоннель.

Как правило, начальные значения (19) являются параметрами стационарного режима, предшествующего нестационарному процессу. В этом случае их можно найти как решение уравнений (2) – (4) для стационарного случая, т.е. при $\partial p / \partial t = 0$, $\partial V / \partial t = 0$, $\partial T / \partial t = 0$, с граничными условиями (9) – (16) при $t = 0$.

Считается, что в начальный момент времени поезда в тоннеле нет.

4 Методика решения

Итак, решение задачи о движении поезда в тоннеле сводится к решению уравнений нестационарной термогидравлики (2) – (4) при выполнении (6) – (8), с граничными условиями (9) – (18) и начальными данными (19). Для решения использовался алгоритм на основе метода характеристик, изложенный в [2].

В более общем случае, тепловой поток q_w вместо приближенной формулы (8) следует определять из совместного решения внутренней термогидравлической задачи и внешней тепловой задачи.

На основе базовой модели возможны различные упрощения, при оценке соответствующих погрешностей. Частным случаем будет стационарная задача, когда все входные параметры задачи неизменны по времени. При незначительном изменении параметров процесса по времени возможно учитывать динамику как последовательную смену стационарных состояний (квазистационарный подход).

При этом следует учесть, что, при данных размерах туннеля, аэродинамическое возмущение будет достигать центра тоннеля примерно за 12 сек, а тепловое – примерно за 3 минуты (в случае движения поезда со скоростью 30 км/час) и/или еще дольше при отсутствии поезда. Поэтому, с высокой степенью вероятности можно предполагать, что будет иметь место нестационарный процесс, близкий к периодическому (при регулярном движении поездов с постоянным интервалом).

Заключение

Система управления режимом вентиляции предназначена для поддержания положительных температур в тоннеле в зимнее время и избежания обледенения. Представленная в работе математическая модель описывает нестационарный вентиляционный режим, возникающий в тоннеле при входе и перемещении по нему поезда.

Указанная математическая модель учитывает термогидравлические параметры потока (давление, скорость, температура), зависимость плотности, вязкости и т.п. параметров воздуха от давления и температуры, теплообмен через стенки тоннеля и др. Поле давления, скоростей и температур насчитывается одновременно как взаимосвязанные параметры потока.

Применение такой модели позволяет выявить узкие места, где возможно появление отрицательных температур и обледенения, подобрать наиболее эффективные режимы управления работой вентиляторов и калориферов.

Заметим, что для достижения наиболее эффективных результатов система регулирования вентиляционного режима тоннеля должна работать не только в режиме автоматического регулирования, т.е. реагировать на прошедшие события, фиксируемые датчиками (изменение температуры, влажности в тех или иных точках тоннеля), но и как система оптимального управления, выбирающая наилучшую стратегию на основе прогноза будущего состояния параметров потока воздуха в тоннеле при тех или иных внешних условиях и режимах работы. Такой прогноз должен опираться на достаточно полную и адекватную модель, в качестве первого шага к созданию которой могут служить результаты настоящей работы.

Список литературы

- [1] Гендлер С.Г., Беспалов С.Е., Соколов В.А., Горшков Ю.М., Юшковский Э.М. Управление тепловым режимом тоннелей в суровых климатических условиях. Транспортное строительство, № 4, 1990, с. 18 – 22.
- [2] Атавин А.А., Лаврентьев В.Л., Тарасевич В.В. Математическое моделирование сложных трубопроводных сетей // Гидравлические цепи. Развитие теории и приложения. Коллективная монография. /Н.Н.Новицкий, Е.В.Сеннова, М.Г.Сухарев и др. - Новосибирск: Наука, Сибирская издательская фирма РАН. - 2000. -273 с. (Раздел 1.1 - стр. 16 - 30.)
- [3] Ривкин С.Л. Термодинамические свойства газов. Справочник. – 4-е изд., перераб. – М.: Энергоатомиздат, 1987. – 288 с., ил.
- [4] Кушнырев В.И., Лебедев В.И., Павленко В.А. Техническая термодинамика и теплопередача. – М.: Стройиздат, 1986. – 464 с., ил.
- [5] Альтшуль А.Д. Гидравлические сопротивления. – М.: Недра, 1982. – 224 с., ил.
- [6] Абрамович Г.Н. К расчету воздушного сопротивления поезда на открытой трассе и в тоннеле. Труды ЦАГИ им. проф. И.Е. Жуковского, вып. 400. – М.: 1939.
- [7] Радченко В.Д. Сопротивление движению вагонов метрополитена. – М.: Трансжелдориздат, 1957, 212 с.
- [8] Кутателадзе С. С. Основы теории теплообмена. – М.: Атомиздат, 1979.
- [9] Лыков А.В. Тепломассообмен: Справочник. – М.: Энергия, 1972.

МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ СПЕЦИФИКАЦИЙ АЛГОРИТМОВ БОРТОВОГО ИНТЕЛЛЕКТА АНТРОПОЦЕНТРИЧЕСКИХ ОБЪЕКТОВ

Б.Е. Федун¹, Т.И. Федун²

¹ФГУП ГосНИИАС
125319, Москва, ул. Викторенко, 7
boris_fed@gosniias.ru
тел: +7 (095) 157-93-49

²МАИ (ТУ)
125871, Москва, ул. Волоколамское шоссе, 4
тел: +7 (095) 158-49-50, факс: +7 (095) 158-29-77

Ключевые слова: модель объекта, уровни управления, структура алгоритма, имитационные модели, деревья развития

Abstract

The model of the object and the its function process structuring for complex man-machine object is offered. The simulation mathematical model structure is described for different global level of management for estimation algorithm object, marketed on digital computers and algorithm to activity object operator.

Введение

При разработке сложных технических объектов уже на ранних стадиях их проектирования необходимо оценивать эффективность определившегося состава бортовых алгоритмов. Рассматриваются взаимодействующие между собой подвижные технические объекты. Множество этих объектов состоит из подмножеств антропоцентрических объектов (например, самолетов) и безэкипажных объектов типов R и BR .

Антропоцентрические объекты (Антр/объект) [1] имеют на своем борту:

- бортовые измерительные устройства (БИзм/Устройства);
- экипаж, осуществляющий управление Антр/объект в соответствии со своими алгоритмами деятельности через диалоговое общение с информационно-управляющим полем (ИУП) кабины Антр/объект;
- БЦВМ-алгоритмы, реализованные в бортовых цифровых вычислительных машинах (БЦВМ);
- бортовые исполнительные устройства (БИсп/Устройства).

Рассматриваемые безэкипажные объекты (БЭО) типов R и BR имеют на борту:

- бортовые измерительные устройства (БИзм/Устройства);
- БЦВМ-алгоритмы, реализованные в бортовых цифровых вычислительных машинах (БЦВМ);
- Бортовые исполнительные устройства (БИсп/Устройства).

Информация об алгоритмах бортового интеллекта (БЦВМ-алгоритмы + алгоритмы деятельности экипажа) каждого Антр/объекту содержится и следующей технической документации:

- техническом задании (ТЗ) на разработку Антр/объект;
- "Логике работы системы "экипаж — бортовая аппаратура"" (далее "Логика работы..."), вербально описывающей замысел конструкторов по семантическому облику Антр/объект;
- "Спецификации алгоритмов бортового интеллекта".

Спецификации алгоритмов бортового интеллекта содержат: информацию о составе БЦВМ-алгоритмов и алгоритмов деятельности экипажа (АДЭ); входную и выходную информацию каждого алгоритма и описание структуры алгоритма. Описание спецификаций позволяет оценивать степень выполнимости требований технического задания (ТЗ) на разработку Антр/объекта и степень выполнимости замысла конструкторов по его семантическому облику, описанному в "Логике работы...".

Информация об алгоритмах бортового интеллекта каждого БЭО содержится в следующих технических документах: техническом задании (ТЗ) на разработку БЭО; спецификации алгоритмов бортового интеллекта (БЦВМ-алгоритмах).

Антр/объект и БЭО находятся в различных состояниях и в различных взаимодействиях между собой, определяемых спецификой рассматриваемой предметной области.

На Антр/объект различают три глобальных уровня управления ГЛУУ [1]:

- уровень определения текущей цели функционирования (уровень целеполагания — первый ГЛУУ);
- уровень определения рационального способа достижения оперативно поставленной цели (второй ГЛУУ);
- уровень реализации определившегося способа достижения оперативно поставленной цели (третий ГЛУУ).

Бортовые алгоритмы, решающие задачи первого и второго ГЛУУ, составляют так называемое *системообразующее ядро* Антр/объект. Совокупность этих алгоритмов (БЦВМ-алгоритмы + АДЭ) называют алгоритмами бортового интеллекта.

Эти же уровни управления различают и на БЭО. Алгоритмы системообразующего ядра таких объектов состоят только из БЦВМ-алгоритмов, причем алгоритмы первого ГЛУУ строятся как аппроксимации речемыслительных алгоритмов возможного (для этого объекта) бортового оператора [2]. БЭО типов *R* и *BR* различаются степенью развитости алгоритмов первого ГЛУУ и мощностью множества типовых ситуаций функционирования БЭО, которые эти алгоритмы охватывают. БЭО типа *R* (примером таких объектов являются боевые управляемые ракеты) имеют "скромные" алгоритмы первого ГЛУУ в отличие от БЭО типа *BR* (пример, проектируемые в США беспилотные боевые самолеты), которые по качеству работы алгоритмов первого ГЛУУ приближаются в простых расчетных условиях функционирования БЭО к соответствующему по назначению Антр/объекту.

Таким образом, возникает *задача оценки эффективности спецификаций алгоритмов, входящих в системообразующее ядро разрабатываемого Антр/объект.*

Необходимость оценки эффективности алгоритмов бортового интеллекта Антр/объекта, начиная с ранних стадий его проектирования (рассматривается стадия, когда есть только ТЗ, "Логика работы..." и спецификации алгоритмов бортового интеллекта), обсуждается на примере наиболее представительного класса Антр/объектов - самолетов [3, 4, 5].

1 Классификация моделей для оценки спецификаций алгоритмов бортового интеллекта

Перед каждым сеансом функционирования Антр/объекта ставится генеральная задача сеанса, под которую готовится как сам объект, так и его экипаж. Генеральная задача представляется через семантическую сеть типовых ситуаций (ТС) функционирования Антр/объекта. Оперативное назначение (на борту Антр/объекта) в процессе сеанса функционирования текущей ТС – задача первого ГЛУУ. На современных Антр/объектах решение задач первого ГЛУУ возлагается только на экипаж. Для этого на информационно-управляющем поле кабины ИУП с помощью БЦВМ-алгоритмов создается информационная модель внешней и внутри-бортовой обстановки (ИМ-ИУП), которая должна обеспечивать экипажу *ситуационную осведомленность* и *ситуационную уверенность*. Для оценки эффективности этой ИМ-

ИУП создается имитационная математическая модель оценки достаточности ситуационной осведомленности и ситуационной уверенности экипажа проектируемого Антр/объекта. Назовем ее *имитационной математической моделью первого* ГЛУУ (ИММ-ПГЛУУ). Для конкретного Антр/объекта такая модель одна. Она должна обеспечивать моделирование работы Антр/объекта во всех возможных (заданных в ТЗ) генеральных задачах функционирования [1].

Для оценки эффективности спецификаций алгоритмов бортового интеллекта, решающих задачи второго ГЛУУ, разрабатываются имитационные математические модели, каждую из которых назовем *имитационной математической моделью второго* ГЛУУ (ИММ-ПГЛУУ). Для каждой ТС функционирования рассматриваемого Антр/объекта создается своя ИММ-ПГЛУУ.

2 Облик ИММ-ПГЛУУ

Модель разрабатывается для имитации работы экипажа с информационно-управляющим полем кабины в квазиреальном времени и имитирует:

- внешнюю обстановку, представляемую в каждый момент времени ситуационным вектором информационные возможности бортовых измерительных устройств, предназначенных для размещения на проектируемом Антр/объекте. Текущая информация, которую в данный момент поставляет на борт Антр/объекта бортовые измерительные устройства, представляется в модели ситуационным вектором $SV(inf)$;
- ИУП кабины проектируемого Антр/объекта для работы с ним реального оператора, имитирующего действия экипажа по назначению текущей ТС. Текущая информация представляется в модели ситуационным вектором $SV(ИУП)$ и предъявляется оператору в форме кадров индикации на ИУП, взятыми из их описания в спецификациях алгоритмов бортового интеллекта.

Назначаемая оператором с ИУП по текущей информации $SV(ИУП)$ типовая ситуация описывается в ИММ-ПГЛУУ интегрально: условия наступления ТС — время ее протекания, исход ТС.

Разработки ИММ-П ГЛУУ находятся на начальных стадиях.

3 Облик ИММ-ПгЛУУ

Модель создается для каждой ТС оцениваемого Антр/объекта. Исходные данные по условиям начала ТС находятся из результатов моделирования, полученных на ИММ-ПГЛУУ.

Различают два типа ИММ-ПгЛУУ: ИММ-ПгЛУУ(R) для оценки алгоритмов бортового интеллекта Антр/объекта в случае отсутствия его взаимодействия с БЭО и ИММ-ПгЛУУ(BR), когда такое взаимодействие возможно.

ИММ-ПгЛУУ (R) не предполагает работу с реальным оператором, имитирующим деятельность экипажа проектируемого Антр/объекта, а ИММ-ПгЛУУ (BR) предполагает.

ИММ-ПгЛУУ (R) имитирует каждый Антр/объект и каждый БЭО из действующих в рассматриваемой ТС объектов. По отношению к каждому действующему объекту все остальные объекты представляют собой внешний мир (см. рисунок), состояние которого для каждого объекта описывается в ИММ-ПгЛУУ(R) значимым для него ситуационным вектором $SV(st)$. На рисунке (где через АО обозначен Антр/объект, а через ГУУ – ГЛУУ), представлен один из действующих в ТС объектов – Антр/объект – и отдельным блоком обозначен внешний для него мир.

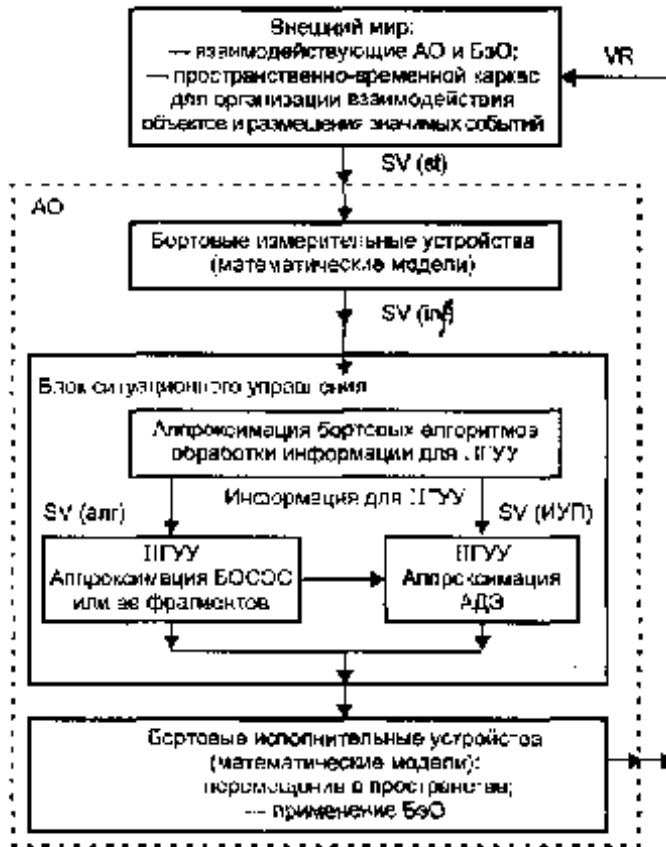


Рисунок 1 – Облик ИММ-ПГУУ(R)/

Блок АО (см. рис.) в своем составе имеет:

- блок математических моделей (ММ) бортовых БИзм/Устройств, размещенных согласно ТЗ на этом Антр/объекте. Через эти устройства на Антр/объекте наблюдается внешний мир, состояние которого на борту Антр/объекта представляется в модели ситуационным вектором $SV(inf)$. В каждый момент времени состав $SV(inf)$ определяется имитируемыми возможностями бортовых БИзм/Устройств и имитируемым качеством работы бортовых алгоритмов первичной обработки информации этого Антр/объекта;
- блок ситуационного управления, в котором имитируются структуры БЦВМ-алгоритмов и АДЭ, решающие задачи ПГЛУУ. Структуры указанных алгоритмов системообразующего ядра рассматриваемого Антр/объекта определяются на основе изучения технических документов "Логика работы..." и "Спецификации алгоритмов бортового интеллекта ТС". В блоке имитируются (аппроксимируются) возможности бортовых алгоритмов обработки информации для задач ПГЛУУ, результаты работы которых представляются через два ситуационных вектора: вектор $SV(алг)$, описывающий доступную в текущий момент на борту Антр/объекта информацию для имитируемых БЦВМ-алгоритмов, и вектор $SV(ИУП)$, описывающий доступную на борту Антр/объекта информацию для имитируемых АДЭ. Оба вектора строятся на основе представляемых в "Спецификациях алгоритмов бортового интеллекта ТС" графов решения оператора (ГРО) и схемах БЦВМ-алгоритмов (СБА) [1]. Блок ситуационного управления имитирует бортовой интеллект Антр/объекта в рассматриваемой ТС;

- блок математических моделей (ММ) бортовых БИсп/Устройств, размещенных согласно ТЗ на этом Антр/объекте. Через эти устройства Антр/объект воздействует на внешний мир. Само воздействие в ИММ-ПГЛУУ (R) описывается вектором VR , текущее состояние которого определяется в модели результатами работы блока ситуационного управления.

Блок БЭО (на рисунке не указан) в своем составе имеет блоки, аналогичные рассмотренным для Антр/объекта:

- блок математических моделей (ММ) бортовых БИзм/Устройств, размещенных согласно ТЗ на этом БЭО;
- блок ситуационного управления, в котором имитируются структуры БЦВМ-алгоритмов, решающие задачи ПГЛУУ этого объекта. Структуры указанных алгоритмов системообразующего ядра рассматриваемого БЭО определяются на основе изучения технических документов по этому объекту.

Блок ситуационного управления имитирует бортовой интеллект БЭО;

- блок математических моделей (ММ) бортовых БИсп/Устройств, размещенных на этом БЭО.

Среди событий ТС (внешнего и внутрибортового мира) выделим события двух типов:

- события, на которые необходимо реагировать (значимые события первого типа);
- события физического контакта Антр/объекта с БЭО типа R, влияющие на факт существования Антр/объекта (значимые события второго типа).

Выделенные события назовем значимыми событиями.

Значимые события второго типа приводят к случайному событию: Антр/объект существует с вероятностью $P_{сущ}$; Антр/объекта не существует с вероятностью $P_{несущ} = 1 - P_{сущ}$. Эти вероятности связываются с той ветвью развития ТС, которая привела к этому значимому событию.

Разработанная для рассматриваемой ТС в "Логике работы..." и в спецификациях бортовых алгоритмов семантическая сеть проблемных субситуаций (ПрС/С) в ИММ-ПГЛУУ (R) отождествляется с семантической сетью значимых событий первого типа. Значимые события второго типа свяжем с достижением антропоцентрического объекта БЭО в конце сеанса его функционирования.

БЭО типа R при определенных условиях появляются в ИММ-ПГЛУУ (R) с задачей (целью) преследования определенного Антр/объекта. При достижении БЭО этого Антр/объекта последний может как прекратить свое существование в ИММ-ПГЛУУ (R), так и продолжить его. Выбор того или иного исхода осуществляется оператором-исследователем, работающим с ИММ-ПГЛУУ (R). Вероятностная интерпретация этого события происходит при анализе результатов моделирования. Сам БЭО после встречи с Антр/объектом или по истечении заданного промежутка времени после его появления в ИММ-ПГЛУУ (R) перестает существовать.

С момента появления БЭО типа R и до момента его встречи с соответствующим Антр/объектом в ИММ-ПГЛУУ (R) прослеживается и фиксируется противодействие Антр/объекта его встрече с преследующим БЭО. Формы и моменты начала противодействия определяются алгоритмами бортового интеллекта, имитируемыми (см. рисунок) в блоке ситуационного управления ИММ-ПГЛУУ (R).

На основе анализа (своя методика для каждого типа объектов) технической документации ТЗ, "Логике работы...", "Спецификации алгоритмов бортового интеллекта" (ГРО&СБА) [1]:

- формируется "множество значимых событий ТС" (семантическая сеть ПСС; бортовые алгоритмы, обеспечивающие идентификацию сложившейся ПСС по $SV(\inf)$ и 5Г(ИУП) как одной из полного множества ПрС/С, представленного в "Логике работы...");
- выделяются из представленных спецификаций БЦВМ-алгоритмов и АДЭ значимые алгоритмы, определяющие полные множества возможных переходов из каждой ПрС/С в другие;
- АДЭ представляются через элементарные акты выработки решения и рассчитываются [2] временные затраты экипажа на принятие каждого решения.

На основе результатов этого анализа разрабатывается блок ситуационного управления ИММ-ПГУУ(R), имитирующий алгоритмы бортового интеллекта рассматриваемого Антр/объект (см. рисунок 1).

Оценка эффективности алгоритмов бортового интеллекта с помощью ИММ-ПГУУ(R). Оператор-исследователь с помощью ИММ-ПГУУ(R) строит дерево развития ТС (ДР ТС), представляющее собой марковскую сеть, узлы которой соответствуют значимым событиям второго типа. ИММ-ПГУУ(R) при наступлении любого значимого события второго типа останавливается и приглашает оператора-исследователя назначить одно из двух возможных состояний Антр/объекта, для которого наступило это значимое событие: существует или не существует. Существование БЗО в ИММ-ПГУУ(R) определяется ситуационно. Оператор-исследователь для себя фиксирует выбранное им состояние Антр/объекта с тем, чтобы при следующей остановке ИММ-ПГУУ(R) на этом же значимом событии, наступающей при повторном воспроизведении (моделировании) всей последовательности предшествующих ему значимых событий, выбрать для этого Антр/объекта другое состояние. Так оператор-исследователь поступает с каждым значимым событием второго типа, работая со всеми Антр/объектами. После прохождения всех возможных последовательностей значимых событий, исходящих из фиксированного начального состояния ТС, оператор-исследователь строит ДР ТС.

При наступлении в ИММ-ПГУУ(R) значимого события первого типа в блоке ситуационного управления соответствующего объекта по текущей информации $VS(\text{алг})$ и $VS(\text{ИУП})$ для Антр/объекта и $VS(\text{алг})$ для БЗО вырабатывается решение реализовавшейся ПрС/С, которое в блоке математических моделей (ММ) бортовых Бисп/Устройств преобразуется в вектор VR (вектор воздействия на внешний мир). Это решение влияет на дальнейший процесс протекания ПрС/С, однозначно определяя его до наступления следующего значимого события.

Работа с ДР ТС проводится оператором-исследователем вне ИММ-ПГУУ(R). Пусть ДР ТС построено. Зададимся начальными значениями вероятностей существования всех рассматриваемых Антр/объектов. На каждой ветви развития ТС по специальному алгоритму, учитывающему особенность соответствующего решения, принятого в ситуационном блоке (см. рис.), и особенности соответствующего БО, рассчитываются вероятности перехода (в построенной марковской сети) каждого Антр/объекта в новое состояние. Исходы ТС – финальные вершины ДР ТС. По марковской сети рассчитываются вероятности наступления каждого возможного исхода ТС. На построенном один раз ДР ТС можно проводить расчеты для различных количественных характеристик $P_{\text{сущ}}$, определяемых особенностями БО и особенностями противодействия им.

Для Антр/объекта "Самолеты" и БО "Ракеты" разработаны ИММ-ПГУУ(R) некоторых ТС, построены ДР ТС и проведены по ним расчеты вероятностей исходов ТС.

Для Антр/объекта «Предприятие» [6] возможно применение такого подхода в случае успешной разработки соответствующей модели этого объекта.

Заключение

Оценка эффективности спецификаций алгоритмов бортового интеллекта сложных Антр/объектов требует разработки двух типов имитационных математических моделей:

- ИММ-ПГУУ для оценки развернутой на ИУП информационной модели внешней и внутри бортовой обстановки с позиций достаточности ее для ситуационной осведомленности экипажа и степени обеспечения ситуационной уверенности экипажа (разрабатывается одна ИММ для рассматриваемого Антр/объект);
- ИММ-ПГУУ для оценки БЦВМ-алгоритмов и АДЭ, разработанных для решения задач ПГУУ (для каждой ТС разрабатывается своя ИММ).

Список литературы

- [1] Федун Б. Е. Конструктивная семантика антропоцентрических систем для разработки и анализа спецификаций бортового интеллекта. М.: Известия РАН, Теория и системы управления (ТиСУ). 1998. № 5.
- [2] Федун Б. Е. Методика экспресс-оценки реализуемости графа решений оператора антропоцентрического объекта на этапе разработки спецификаций алгоритмов бортового интеллекта. М.: Известия РАН. Теория и системы управления (ТиСУ). 2002. №1.
- [3] Авиация ПВО и научно-технический прогресс / Под ред. Е.А. Федосова. М: Дрофа, 2001.
- [4] Волин С. Тактическая авиация: долгий путь к пятому поколению. М.: Вестник авиации и космонавтики. 2002. № 3.
- [5] Федун Б. Е. Бортовые оперативно-советующие экспертные системы тактических самолетов пятого поколения (обзор по материалам зарубежной печати). М.: НИЦ ГосНИИАС, 2002.
- [6] Мавроди С.И., Федунова Т.И. и др. Управление персоналом предприятия. Отчет МАИ ТУ. 2003.

ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ УПРАВЛЕНИЯ ПРИ АВАРИЙНОМ ВЫБРОСЕ РАДИОАКТИВНЫХ ВЕЩЕСТВ В АТМОСФЕРУ

Н.Ю. Истомина, А.Д. Истомин, А.Н. Жиганов, М.Д. Носков

Северский государственный технологический институт
636036, Северск, пр. Коммунистический, 65, Россия
E-mail istomin@ssti.ru
Тел./факс (38242) 2-93-88.

***Ключевые слова:** базы знаний, геоинформационные системы, моделирование, прогнозирование, ликвидация последствий, экспертные системы, реципиенты риска, потенциально-опасные объекты*

Abstract

Concept, structure and functions of geoinformation expert modeling complex are presented. The complex is designed for postaccident assessment, and emergency planning in the case of accidental release of radioactive substances into the atmosphere. The application programme created on the base of the proposed concept is presented.

Введение

Возникновение аварийной ситуации, связанной с выбросом радиоактивных веществ в атмосферу, требует оперативного принятия решения по снижению воздействия ионизирующего излучения на население и окружающую среду. Решение по управлению ситуацией, сложившейся в результате выброса, принимается на основе информации о радиационной обстановке на местности, формирование которой зависит от множества разнородных, меняющихся с течением времени факторов. Неадекватность решений может быть вызвана несвоевременностью, недостатком, противоречивостью информации, а также ошибками при обработке и анализе значительных объемов разнородных данных. В связи с этим при принятии управленческих решений целесообразно использовать специализированные программно-технические средства, созданные на основе современных информационных технологий. В статье представлены концепция геоинформационного экспертно-моделирующего комплекса (ГИЭМК) и описание разработанного на ее основе проблемно-ориентированного программного обеспечения, предназначенного для поддержки принятия решений по минимизации последствий аварийного выброса радиоактивных веществ в атмосферу.

1 Геоинформационная система

Пространственный характер данных мониторинга, а также необходимость представления результатов с привязкой к карте местности требует использования современных геоинформационных технологий [1-3]. Геоинформационная система комплекса предназначена для отображения пространственно-временных и атрибутивных данных на карте рассматриваемой территории. Визуальное представление информации базируется на цифровой модели местности (ЦММ), организованной в виде множества слоев (карт-подложек). Одним из основных слоев геоинформационной системы является физическая карта местности в растровом формате. Реальные объекты, расположенные на рассматриваемом участке территории, отображаются в векторном формате и образуют другие слои ЦММ. Каждому реальному объекту сопоставляется объект геоинформационной системы, который характеризуется набором пространственных и тематических данных. Для определения положения объектов на карте местности использует-

ся декартова система координат. Тематически реальные объекты разделяются на объекты – реципиенты техногенного риска и потенциально-опасные объекты. Как реципиенты техногенного риска, так и потенциально-опасные объекты геометрически классифицируются как точечные, линейные, площадные.

Точечные объекты реципиенты – контрольные точки. С помощью контрольных точек осуществляется наблюдение временной динамики физических величин в данном положении на местности. Атрибутами контрольных точек являются их координаты, условный знак и название физической величины.

Линейными реципиентами риска являются транспортные коммуникации, реки. Линейные объекты, также как и точечные, имеют свои атрибуты. Линейные объекты представляются полилинией, состоящей из упорядоченного набора связанных отрезков (звеньев), с координатными данными и атрибутами: узловые точки звена, задаваемые парой координат (x, y) , параметры (цвет, толщина) линии, отображающей звено, направление движения, интенсивность движения (для транспортных коммуникаций), список названий и значений физических величин, позволяющих оценить степень воздействия радиоактивных веществ.

Площадные объекты реципиенты изображаются ареалами – областями, ограниченными замкнутой полилинией, задаваемой списком координат всех звеньев. К площадным объектам реципиентам риска относятся населенные пункты. Для населенных пунктов задаются параметры отображения ареала (цвет заливки или тип штриховки), численность населения, возрастной состав, характеристика застройки, средства эвакуации и т.д.

Потенциально-опасными объектами считаются объекты промышленности, функционирование которых сопряжено с риском выброса радиоактивных веществ в атмосферу. Принципы отображения потенциально-опасных объектов те же, что и для реципиентов риска. Каждый потенциально-опасный объект имеет следующие атрибутивные характеристики: высота выброса, тип выброса (непрерывный, кратковременный, мгновенный), время начала выброса (для кратковременного дополнительно – время окончания выброса), суммарная активность выброса – для точечного источника, активность, приходящаяся на единицу длины (площади) – для линейного (площадного) объекта, радионуклидный состав выброса.

2 Моделирующая система

Моделирующая система содержит модуль расчета распространения радиоактивных веществ и модуль расчета воздействия радионуклидов [3, 4]. Модуль для расчета распространения радиоактивных веществ предназначен для определения удельных активностей в воздухе и на подстилающей поверхности. Воздействие ионизирующего излучения определяется с помощью расчета эквивалентных доз. Работа каждого из модулей базируется на соответствующей математической модели. Для описания распространения радиоактивных веществ в приземном слое атмосферы используется метод случайно-блуждающих точек [4]. Математическая модель оценки воздействия радиоактивных веществ базируется на методике расчета эквивалентных доз, обусловленных облаком радиоактивных веществ, загрязненной поверхностью, ингаляцией радионуклидов. Расчеты доз осуществляются на основе данных моделирования распространения радиоактивных примесей с учетом соответствующих характеристик излучения отдельных радионуклидов [5, 6]. Результаты расчетов отображаются на карте местности в виде изолиний поверхностных концентраций, временных графиков и изолиний мощностей эквивалентных доз.

3 Экспертно-аналитическая система

Основная задача экспертно-аналитической системы комплекса экологического мониторинга заключается в анализе сложившейся радиационной обстановки и подготовке рекомендаций с целью минимизации воздействия ионизирующего излучения на население. Экспертно-аналитическая система содержит модули идентификации параметров моделирования, анализа

радиационной обстановки, а также базы данных параметров моделирования распространения примесей, параметров расчета дозовых нагрузок, нормативных документов и аварийных планов.

Модуль идентификации параметров моделирования предназначен для определения значений параметров, необходимых для расчетов распространения облака радиоактивных веществ и дозовых нагрузок. Идентификация параметров проводится согласно логическим правилам, на основе введенных метеоданных и данных ГИС о потенциально-опасных объектах и соответствующих баз данных.

Модуль анализа радиационной обстановки предназначен для анализа сложившейся ситуации и выработки рекомендаций действий по снижению негативного воздействия радиоактивных веществ на население. Анализ сложившейся радиационной обстановки проводится с помощью сопоставления результатов расчетов с критериями безопасности, введенными пользователем и/или содержащимися в базе данных нормативных документов. На основе результатов анализа радиационной обстановки и базы данных аварийных планов вырабатываются рекомендации по проведению мероприятий, направленных на минимизацию последствий выбросов радиоактивных веществ в атмосферу.

Работа экспертно-аналитической системы включает в себя несколько этапов. На первом этапе, на основе введенных пользователем характеристик выброса (высоты, геометрической формы, интенсивности выброса, химического состава выброса), метеоданных (скорости и направления ветра, облачности, инсоляции, интенсивности осадков) определяется модель распространения радиоактивных веществ и ее параметры. Диалоги ввода и редактирования данных представлены на рисунке 1. На втором этапе, на основе результатов расчетов, базы данных о населенных пунктах, средствах, маршрутах эвакуации, с учетом действующих критериев,

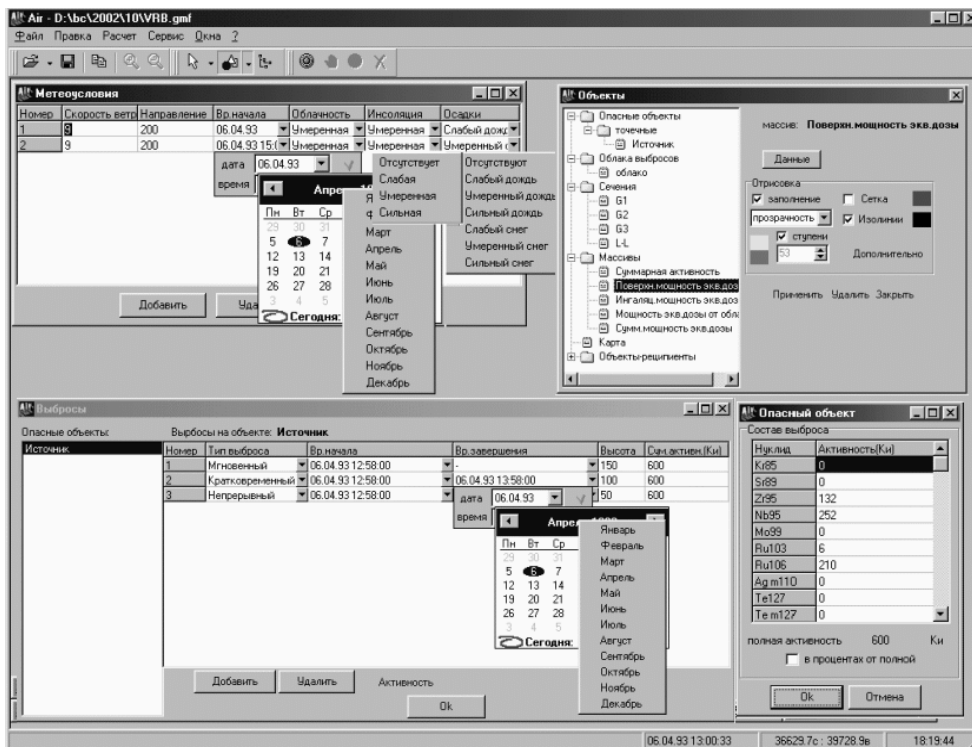


Рисунок 1 – Диалоги ввода и редактирования данных программного комплекса

правил и предписаний [6] осуществляется поиск решений согласно приоритетности рекомендуемых мер. На третьем этапе происходит непосредственное представление решения ответственному лицу. После корректировки решения в соответствии с нормативной базой данных происходит его документирование

4 Проблемно-ориентированное программное обеспечение

Представленная концепция комплекса экологического мониторинга была реализована в виде проблемно-ориентированного программного обеспечения, представляющего собой многопоточное, многооконное приложение, предназначенное для работы на персональном компьютере в операционной системе Windows 98 и выше. При запуске программного обеспечения ядро операционной системы создает для него процесс и выделяет некоторое изолированное адресное пространство. При запуске процесса создается главный управляющий поток приложения. Запуск программы активизирует главное окно программы, содержащее заголовок, панель меню, панель инструментов, рабочую область, статус строку. Главное окно приложения предназначено для работы с проектом, представляющим собой совокупность пространственно – атрибутивной информации о потенциально-опасных объектах и реципиентах радиационного риска, пространственно – временных данных распределения физических величин. Проект не обладает какими-либо инструментами для манипулирования собственными данными. Ввод и редактирование данных, управление расчетом и представлением результатов, сохранение данных осуществляется с помощью блоков: управления проектом, управления объектами геоинформационной системы и параметрами моделирования, визуализации данных, управления расчетом, обработки команд.

Блок управления проектом, предназначен для создания нового проекта, открытия существующего проекта из базы данных, сохранения текущего проекта, сохранения его под другим именем, завершения работы с проектом. Блок управления объектами геоинформационной системы и параметрами моделирования предназначен для ввода и редактирования данных объектов, включая месторасположение, высоту и тип источника выброса; радиоизотопный состав и активность выброса; метеоданные (скорость и направление ветра, степень инсоляции, облачность, интенсивность и вид осадков) для различных интервалов времени, редактирования и удаления комбинаций метеоусловий. Блок управления расчетом позволяет запускать процесс моделирования, приостанавливать расчет, останавливать расчет, прерывать процесс расчета, обеспечивать доступ к данным блоку визуализации данных. Блок визуализации данных предназначен для отображения на карте местности потенциально-опасных объектов и объектов реципиентов, реперов, сечений, изолиний, результатов работы моделирующей и экспертно-аналитической систем, а также для управления режимом отображения данных в главном и дочерних окнах. На рисунке 2 представлены некоторые из возможностей представления результатов работы ГИЭМК. Результаты моделирования представляются в виде изолиний массивов: концентрации радионуклидов на поверхности, мощности эквивалентной дозы на поверхности, мощности эквивалентной дозы от облака, ингаляционной мощности дозы и суммарной мощности дозы, а также в виде графических зависимостей указанных величин от времени и по сечению. Результатами работы экспертно-аналитической системы являются: решения о мерах защиты населения объектов реципиентов; зонирование территории, подвергшейся радиоактивному загрязнению; оптимальный аварийный план. Информация о радиационной обстановке и мерах защиты населения предоставляется пользователю в виде таблиц и карт зонирования местности. Экспертно-аналитическая система также предоставляет пользователю возможность формирования запросов на поиск объектов с заданными характеристиками (например, уровень активности и/или мощности дозы на поверхности), введения контрольных точек, сечений, участков поверхности с целью получения интегральных значений параметров этих объектов.

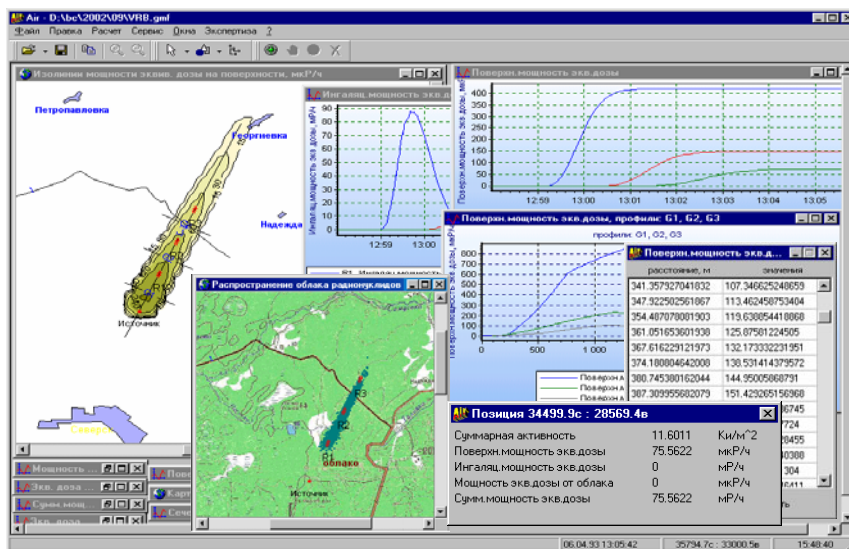


Рисунок 2 – Возможности представления результатов работы ГИЭМК.

Заключение

Областями представленного в работе Геоинформационно-моделирующего комплекса являются: обработка данных мониторинга загрязнения радиоактивными веществами подстилающей поверхности и воздушного бассейна; моделирование распространения радиоактивных примесей в атмосфере вследствие аварийных ситуаций или штатной работы предприятия, определение удельных активностей на поверхности и в атмосфере; проведение расчетов эквивалентных доз, обусловленных загрязненной поверхностью, облаком радиоактивных примесей, ингаляцией радионуклидов, а также коллективных доз населения; разработка рекомендаций по ликвидации последствий выбросов радиоактивных веществ в приземный слой атмосферы.

Описанный комплекс является достаточно универсальным. Разработанное проблемно-ориентированное программное обеспечение может быть использовано для создания систем экологического мониторинга различного назначения.

Список литературы

- [1] Геловани В.А., Башлыков А.А., Бритков В.Б. Интеллектуальные системы поддержки принятия решений в нештатных ситуациях с использованием информации о состоянии природной среды. – М.: Эдиториал УРСС, 2001. – 304 с.
- [2] Цветков В.Я. Геоинформационные системы и технологии. – М.: Финансы и статистика, 1998. – 288 с.
- [3] Истомина Н.Ю., Истомина А.Д., Носков М.Д. Применение ГИС для прогнозирования распространения загрязняющих веществ и оценки их воздействия на человека // Сб. научных трудов Национальной горной академии Украины. Т. 1. № 9. – Днепропетровск: РИК НГА Украины, 2000. С. 164–168.
- [4] Жиганов А.Н., Истомина Н.Ю., Носков М.Д. Моделирование последствий выброса радиоактивных веществ в атмосферу // Известия вузов. Физика. 2000. Т. 43, № 4 (приложение). С. 100-104.
- [5] Гусев Н.Г., Беляев В.А. Радиоактивные выбросы в биосфере: Справочник. – М.: Энергоатомиздат, 1986. – 224 с.
- [6] Нормы радиационной безопасности (НРБ-99): Гигиенические нормативы. – М.: Центр санитарно-эпидемиологического нормирования, гигиенической сертификации и экспертизы Минздрава России, 1999. – 116 с.

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ УПРАВЛЕНИЯ РАЗРАБОТКОЙ МЕСТОРОЖДЕНИЙ УРАНА МЕТОДОМ ПОДЗЕМНОГО ВЫЩЕЛАЧИВАНИЯ

А.Г. Кеслер, А.Н. Жиганов, А.Д. Истомин, М.Д. Носков

Северский государственный технологический институт

636070, Северск, пр. Коммунистический, 65, Россия

kesler@ssti.ru

тел: +7 (242) 4-93-88

Ключевые слова: управление разработкой, математическое моделирование, геотехнологическая информационная система, подземное выщелачивание урана

Abstract

The program complex for the management of uranium field development by in-situ leaching method is presented. The program complex consist of geotechnological and modeling systems. The geotechnological system allows input, edit, extrapolation, and visualization of the geological, hydrogeological and mineralogical parameters of ore body. The modeling system describes hydrodynamical and physical-chemical processes taking place in underground water-bearing strata. The program complex can be used for increase of field development profitability, planning of mine working, and reduction of the environment pollution.

Введение

Одним из перспективных способов добычи полезных ископаемых является подземное выщелачивание (ПВ) [1]. Методом ПВ добываются такие металлы как уран, золото, медь, никель, редкоземельные элементы и т.д. С помощью ПВ разрабатываются экзогенные месторождения полезных ископаемых, которые образуются в хорошо проницаемых водоносных горизонтах. Извлечение полезных компонентов из рудного тела происходит с помощью системы технологических скважин. Через нагнетательные скважины в продуктивный горизонт закачивается раствор веществ, способных вступать в химические реакции с минералами, содержащими полезные компоненты. В результате происходит переход полезных компонентов в жидкую фазу, после чего они выводятся из продуктивного горизонта через откачные скважины.

Технологические и экономические параметры разработки месторождений методом ПВ зависят от геологического строения водоносного комплекса, состава рудовмещающих пород, гидрогеологических и физико-химических процессов, происходящих в рудосодержащих пластах, особенностей фильтрации растворов и т.д. Ввиду высокой сложности и взаимосвязанности процессов, происходящих в подземных водоносных горизонтах при добыче полезных ископаемых методом ПВ для совершенствования технологии подземного выщелачивания и методов управления процессом, целесообразно использовать современные информационные технологии [1, 2].

1 Описание программного комплекса

В данной работе рассматривается программный комплекс, предназначенный для проектирования и управления разработкой месторождений урана методом ПВ. Комплекс включает в себя геотехнологическую информационную и моделирующую системы. Геотехнологическая информационная система позволяет вводить, редактировать, экстраполировать и представлять с привязкой к карте местности полный набор исходных данных, необходимых для моделирования изменения состояния продуктивного горизонта в процессе ПВ. Она состоит из четырех

функциональных блоков, описывающих геологические, гидрогеологические, минералогическо-геохимические параметры продуктивного горизонта и технологические характеристики эксплуатационных скважин. Геологический блок включает в себя привязанные к карте местности данные о геометрии участка водоносного пласта, в котором находится рудное тело (глубина залегания, мощность продуктивного горизонта, расположение рудного тела и т.д.). Гидрогеологический блок содержит данные о распределении фильтрационных параметров продуктивного горизонта, а так же другие данные, необходимые для описания миграции жидкости в пористой среде. Минералогическо-геохимический блок содержит данные о начальном распределении минералов в продуктивном горизонте и компонентов, растворенных в пластовых водах. Также в него включены кинетические коэффициенты и параметры равновесия физико-химических процессов, протекающих в системе. Технологический блок содержит данные о положении и вскрытии скважин, режимы их работы, составы закачиваемых растворов.

Моделирующая система основывается на комплексной математической модели многокомпонентной фильтрации [2] и состоит из двух блоков, описывающих гидродинамические и физико-химические процессы. Гидродинамический блок включает в себя расчет распределения давления и фильтрационных потоков в приближении жесткого режима фильтрации. Гидродинамические расчеты выполняются с учетом гидродисперсии, режимов работы технологических скважин, неоднородности фильтрационных параметров продуктивного горизонта и регионального потока. В химический блок включены расчеты растворения урансодержащих минералов кислотой, потери кислоты в результате взаимодействия с кислотопоглощающими минералами, переотложение минералов, комплексообразование, сорбция и десорбция. При расчетах учитывается неоднородность минералогического строения породы, кинетика взаимодействия технологического раствора с различными минералами, неоднородность состава подземных вод и переменный состав растворов, закачиваемых в продуктивный горизонт.

На основе описанной модели был создан программный комплекс разработки месторождений урана методом ПВ предназначенный для работы на персональном компьютере (рисунок 1). Комплекс создан на языке Borland Builder C++ и представляет собой проблемно-ориентированное многопоточное, многооконное, 32-битное программное обеспечение, работающее под управлением операционной системы Windows 98–XP. Использование средств визуального программирования при разработке программного комплекса предоставляет пользователю возможность работы со стандартными интерфейсными объектами – окнами, списками, таблицами, которые легко связываются с базами данных и отображаются на экране монитора. Программный комплекс обладает также стандартными элементами управления – кнопки, переключатели, флажки, меню, всплывающие подсказки на русском языке. Запуск приложения активизирует главное окно, содержащее заголовок, панель меню, панель инструментов, рабочую область, статус строку, отображающую контекстные подсказки, расчетное время, координаты положения курсора на карте, текущее системное время. Главное окно приложения предназначено для работы с проектом, представляющим собой некоторую совокупность пространственно-атрибутивной информации о полигоне ПВ, геологическом, минералогическом строении и временной эволюции состояния продуктивного горизонта. Ввод и редактирование данных, управление расчетом и представлением результатов, сохранение данных осуществляются с помощью блоков управления проектом, управления объектами геоинформационной системы и параметрами моделирования, визуализации данных, управления расчетом, обработки команд.

Блок управления объектами геоинформационной системы и параметрами моделирования предназначен для ввода и редактирования геологических, гидрогеохимических, минералогических и технологических данных о продуктивном горизонте, положении и режимах работы скважин и составах закачиваемых растворов. Кроме этого возможно редактирование параметров математической модели процесса ПВ. Блок управления расчетом позволяет: запускать процесс моделирования, приостанавливать расчет, останавливать расчет, прерывать процесс

расчета, а также обеспечивает передачу данных в блок визуализации. Блок визуализации данных предназначен для отображения на карте местности скважин, реперов, сечений, изолиний, а также для управления режимом отображения данных в главном и дочерних окнах. Результаты моделирования представляются блоком визуализации данных в виде изолиний и/или картин пространственного распределения исходных и текущих концентраций компонентов в жидкой и твердой фазах, содержания минералов, параметров состояния продуктивного горизонта (пористость, мощность и др.), профилей этих величин вдоль задаваемых пользователем сечений. Кроме этого, возможен вывод временных зависимостей значений физических величин в заданных точках (скважинах или реперах), а также технологических показателей по скважинам и по всему полигону.

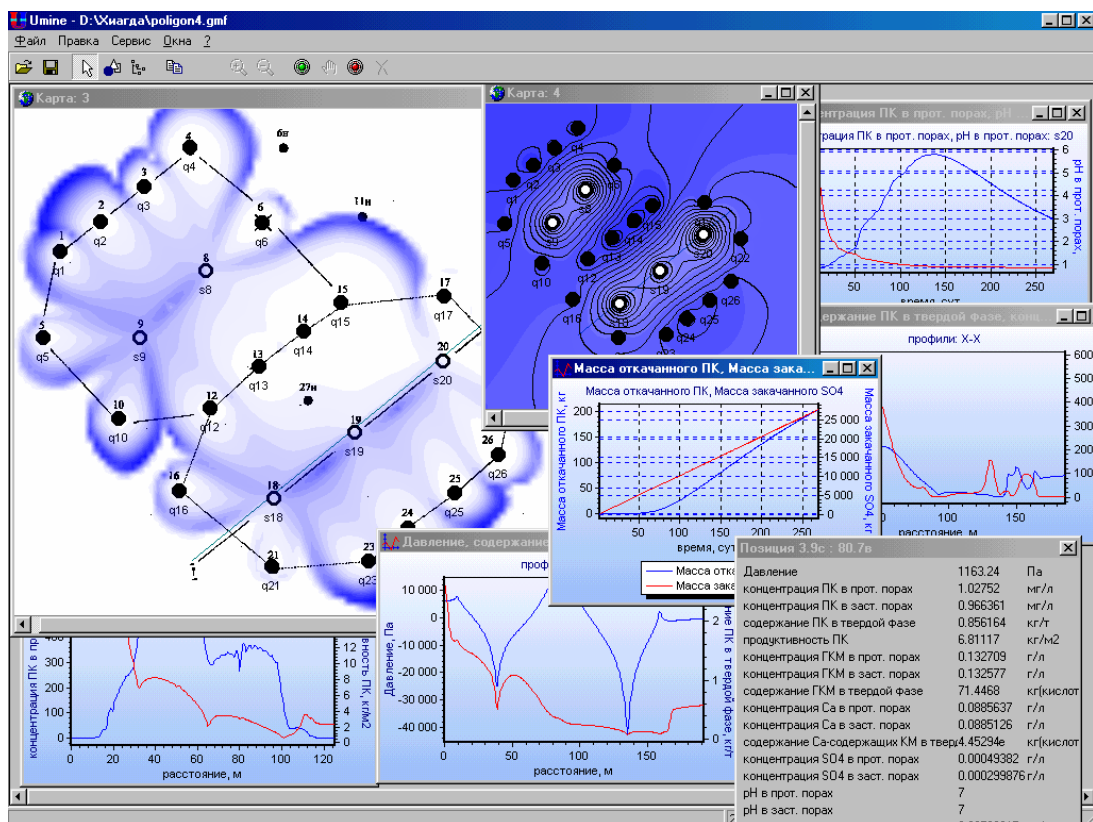


Рисунок1 – Вид рабочего окна программного комплекса для управления разработкой месторождений урана методом ПИВ

2 Последовательность применения программного комплекса

Применение программного комплекса осуществляется в несколько этапов. Первым этапом использования комплекса является восстановление исходного состояния продуктивного горизонта в области моделирования. Для этого загружается растровая карта или план технологической площадки. Затем, с помощью системы диалогов задается система координат, вводятся положение и параметры технологических и наблюдательных скважин. Создание цифровой модели, описывающей гидрогеохимическое, минералогическое и геологическое строения продуктивного горизонта, может производиться двумя способами. Первый способ – непосредствен-

ный ввод и редактирование таблиц, описывающих распределения гидрогеохимических, минералогических и геологических параметров продуктивного горизонта. С помощью системы диалогов можно задать или изменить (прибавить, умножить, отнять поделить на заданное число) значение выбранной физической величины сразу во всей таблице, в произвольной прямоугольной области или в одной ячейке. Другой метод основан на интерполяции данных. Пользователь задает на карте точечные объекты (реперы) или ломанную линию (изолинию, замкнутую или незамкнутую). После этого определяет значение одной или нескольких величин в этой точке или вдоль изолинии. На основе введенной информации геоинформационная система восстанавливает данные во всем продуктивном горизонте с помощью заданной процедуры аппроксимации.

Следующим этапом является определение параметров физико-химической модели. Идентификация параметров может осуществляться на основе экспериментов по фильтрации технологических растворов через колонки, содержащие реальные образцы вмещающих пород продуктивного горизонта, а также сравнения результатов компьютерного моделирования и данных двухскважинного или односкважинного опыта по выщелачиванию урана из рудовмещающего пласта. Цифровая модель продуктивного горизонта, содержащая технологическую, гидрогеохимическую, минералогическую, геологическую и физико-химическую информацию, может быть использована для оценки запасов урана, визуализации продуктивного горизонта и рудного тела, а также в качестве начальных данных для моделирования добычи урана методом подземного выщелачивания из данного продуктивного горизонта.

На третьем этапе производится моделирование эксплуатации блока рассматриваемого месторождения и верификация математической модели. В качестве исходных данных используются цифровая модель продуктивного горизонта, режимы работы нагнетательных и откачных скважин, а так же составы нагнетаемых растворов. Расчет динамики процесса ПВ осуществляется с параметрами моделирования, полученными на основе лабораторных экспериментов и двухскважинных (односкважинных) опытов. Результаты моделирования сравниваются с реальными данными разработки блока месторождения. В случае несовпадения расчетных и натуральных данных происходит уточнение параметров моделирования и производится повторный расчет.

На основе уточненных параметров моделирования и цифровой модели продуктивного горизонта проводятся исследования различных схем разработки блока месторождения. Осуществляется поиск наиболее выгодных с точки зрения увеличения доли извлеченного урана и снижения потерь серной кислоты режимов работы технологических ячеек. При этом варьируются объемы и составы нагнетаемых растворов, объемы откачиваемых растворов, исследуется влияние противofильтрационных завес на структуру массопотоков в продуктивном горизонте, влияние периодов простоя на характер извлечения урана и т.д. Так же рассматривается целесообразность ввода новых скважин в эксплуатацию и поиск оптимальных схем размещения технологических скважин в разрабатываемом блоке.

В качестве примера применения геотехнологического информационно-моделирующего комплекса, было проведено моделирование разработки блока 8(9-12) Далматского месторождения урана, Россия. Блок включал в себя три откачных и восемь закачных скважин. На основе натуральных данных о минералогическом, геологическом и гидрогеологическом строении продуктивного горизонта в скважинах, с помощью аппроксимации была создана цифровая модель ячейки месторождения. Восстановленное распределение концентрации урана в твердой фазе приведено на рисунке 2а. Физико-химические параметры определялись на основе сравнения результатов компьютерного моделирования и данных двухскважинного опыта. При этом использовались реальные характеристики скважин и режимы их работы. На рисунке 3 приведено сравнение интегральных технологических показателей (массы откаченного урана и серной кислоты), полученных в ходе опытной разработки блока Далматского месторождения, и компьютерного моделирования. Распределение концентрации урана в твердой фазе спустя четыре

года эксплуатации блока приведено на рисунке 2б. Распределения концентрации урана в жидкой фазе и водородного показателя pH в продуктивном горизонте на момент окончания эксплуатации приведены на рисунке 4. Хорошее соответствие расчетных и экспериментальных результатов подтверждает адекватность модели и правильность расчетов разработки месторождения методом ПВ.

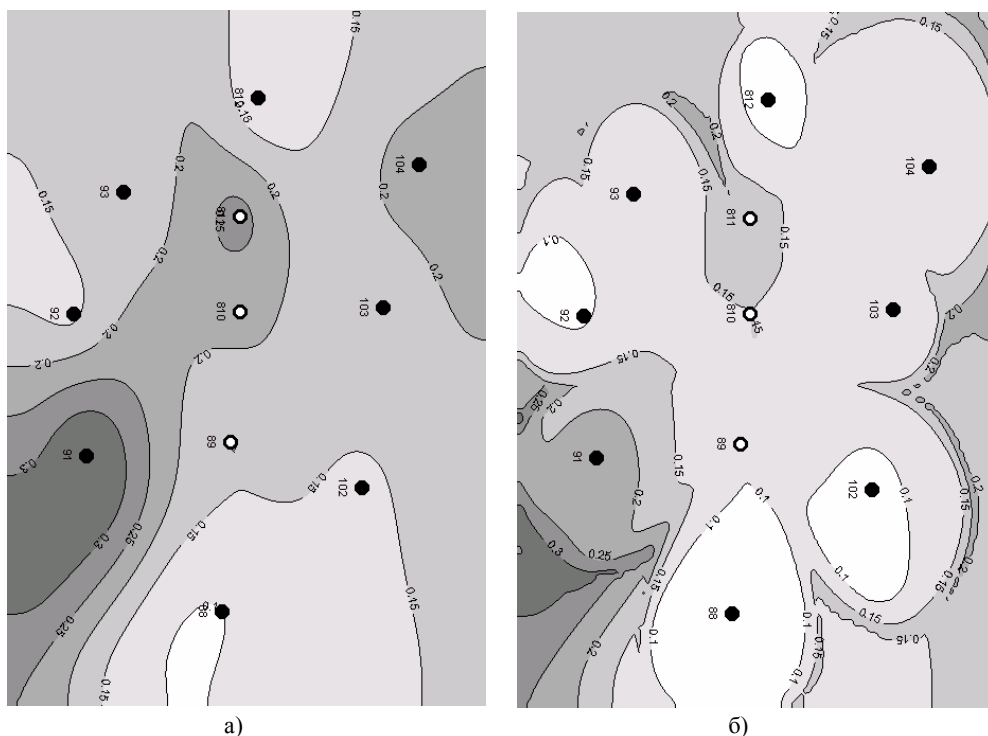


Рисунок 2 – Распределение концентрации урана в твердой фазе (кг_{урана}/Т_{породы}): а) на момент начала разработки; б) через четыре года после начала разработки блока.

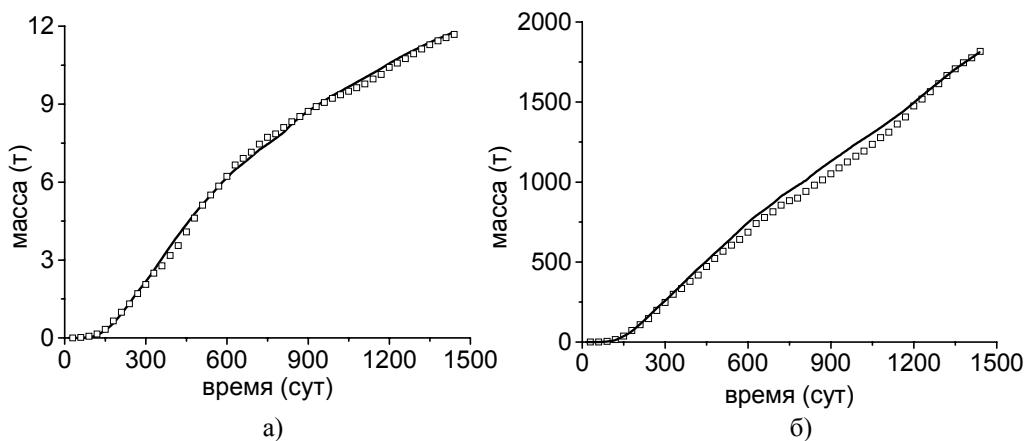


Рисунок 3 – Зависимость: а) массы извлеченного урана; б) массы отквашенной кислоты от времени. □ – экспериментальные данные, — – результаты моделирования.

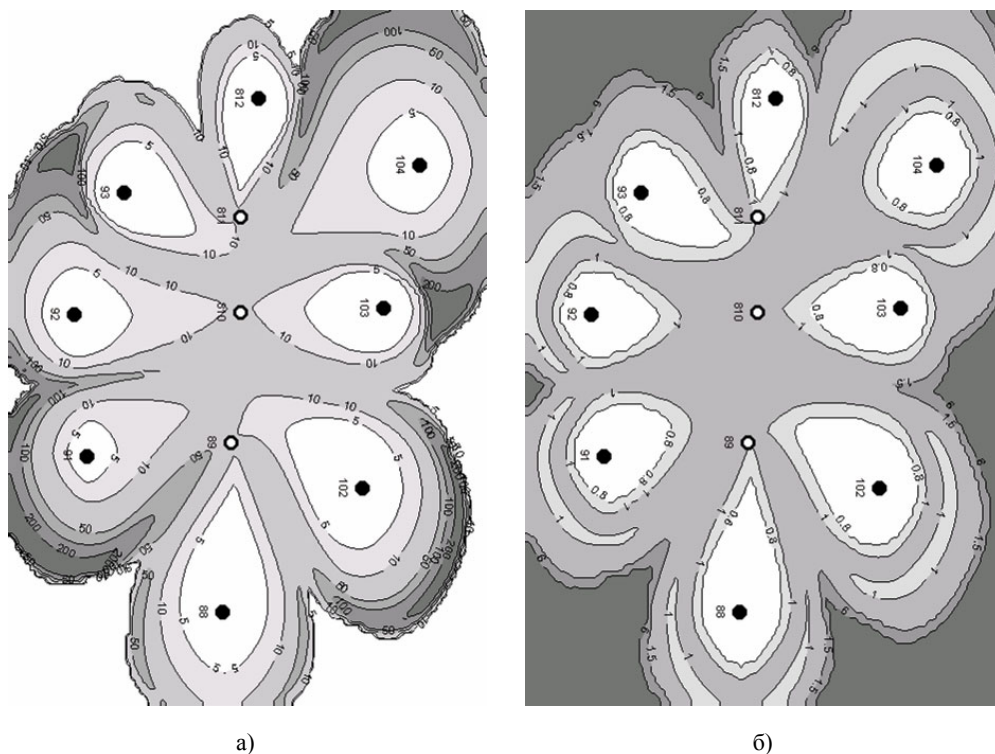


Рисунок 4 – Распределение а) концентрации урана в жидкой фазе (мг/л) и б) pH через четыре года после начала разработки блока.

Заключение

Программный комплекс, позволяет моделировать разработку месторождения урана методом подземного выщелачивания. Используемая в комплексе математическая модель адекватно описывает физико-химические процессы в продуктивном горизонте при сернокислотном выщелачивании урана. Программный комплекс может быть использован при проектировании и разработке месторождений урана методом ПВ для увеличения доли извлеченного урана, уменьшения расходов на единицу продукции и минимизации загрязнения подземных вод. На существующих месторождениях, разрабатываемых с помощью ПВ, его можно использовать для обоснования ввода в эксплуатацию новых скважин, подбора режима эксплуатации скважин, обеспечивающего максимальное извлечение урана при минимальных затратах. На разведанных месторождениях программный комплекс может быть использован для: определения оптимального расположения нагнетательных и откачных скважин, подбора режимов эксплуатации скважин, обеспечивающих максимальную эффективность разработки.

Список литературы

- [1] Лаверов Н.П., Абдульманов И.Г., Бровин К.Г., и др. Подземное выщелачивание полиэлементных руд. – М.: Издательство академии горных наук, 1998. – 448 с.
- [2] Жиганов А.Н., Носков М.Д., Истомина А.Д., Кеслер А.Г., Чеглоков А.А. Построение геотехнологической информационно-моделирующей системы для управления разработкой месторождения методом подземного выщелачивания // Проблемы управления и моделирования в сложных системах: Труды III Международной конференции – Самара: Самарский научный центр РАН, 2001. С. 588–593.

ОПТИМАЛЬНЫЙ ПО БЫСТРОДЕЙСТВИЮ РЕГУЛЯТОР ДЛЯ ПРОИЗВОЛЬНОГО ЛИНЕЙНОГО ОБЪЕКТА ВТОРОГО ПОРЯДКА

Ю.Н. Золотухин, А.А. Нестеров

Институт автоматки и электрометрии СО РАН
630090, Новосибирск, пр. Коптюга, 1
zol@idisys.iae.nsk.su
тел: +7 (3832) 33-26-25, факс: +7 (3832) 33-38-63

Ключевые слова: оптимальное управление, объект второго порядка, моделирование

Abstract

This paper presents an optimal control problem solution for an arbitrary linear second order plants.

Введение

Несмотря на создание общей теории оптимальных систем, задача синтеза [1] оптимального по быстродействию управления решена только для некоторых частных случаев линейных объектов второго или третьего порядка.

Постановка задачи и решение

Ниже описан регулятор, осуществляющий синтез оптимального по быстродействию управления для объекта, описываемого уравнением

$$(1) \quad \dot{X} = AX + BU,$$

где:

- X - двумерный вектор координат объекта;
- U - двумерный в общем случае вектор управляющих параметров, ограниченных условиями

$$(2) \quad |U_i| \leq 1, i = 1, 2;$$

- A и B - произвольные матрицы, ограниченные только требованиями общности положения, т.е. векторы BW , ABW - линейно независимы, W - вектор, параллельный любому из ребер квадрата (2).

В соответствии с принципом максимума [1] вектор $U(t)$, переводящий систему (1) из положения $X(t_0) = X_0$ в положение $X(T) = 0$ за минимальное время T , определяется соотношением

$$(3) \quad U(t) = \text{sign}(B^T \Psi(t)),$$

где B^T - транспонированная матрица B , а вектор $\Psi(t)$ определяется уравнением

$$(4) \quad \dot{\Psi}(t) = -A^T \Psi.$$

В этом случае решение системы (1) запишется в виде

$$X(t) = e^{At} \left[X_0 + \int_0^t e^{-A\tau} B \text{sign}(B^T \Psi(\tau)) d\tau \right]$$

или, с учетом граничных условий,

$$(5) \int_0^T e^{-A\tau} B \operatorname{sign}(B^T \Psi(\tau)) d\tau = -X_0.$$

Из уравнения (4) получаем

$$\Psi(t) = e^{-A^T t} \Psi_0,$$

где $\Psi_0 = \Psi(0)$, и, следовательно,

$$(6) B^T \Psi(t) = B^T e^{-A^T t} \Psi_0 = (e^{-At} B)^T \Psi_0.$$

Введем матрицу

$$(7) V(t) = e^{-At} B,$$

которая является решением уравнения

$$(8) \dot{V} = -AV, V(0) = B,$$

и преобразуем (5) с учетом (6) и (7):

$$(9) \int_0^T V(t) \operatorname{sign}(V^T(t) \Psi_0) dt = -X_0.$$

Уравнение (9) связывает начальные условия сопряженной системы Ψ_0 с координатами управляемой системы. Если рассматривать точку X_0 как текущую точку $X(t)$ решение уравнения (9) должно позволить определить Ψ_0 как функцию текущей координаты $X(t)$. В этом случае мы можем решить задачу синтеза в виде

$$(10) U(X) = \operatorname{sign}(B^T \Psi(X)).$$

В работах [2] и [3] были исследованы некоторые свойства уравнения (9) и предложен способ получения соотношений (10).

Прежде всего, преобразуем матрицу B к виду $\begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$ для системы с двумя управляющими параметрами и к виду $\begin{bmatrix} 0 \\ b \end{bmatrix}$ для одного управляющего параметра. Такое преобразование всегда возможно и является обратимым.

Ниже представлена процедура синтеза оптимального управления в соответствии с [2, 3].

Для одного управляющего параметра $V(t) = \begin{bmatrix} V_1(t) \\ V_2(t) \end{bmatrix}$ и задача синтеза решается уравнениями

$$(11) \begin{cases} \int_0^T |V_1(t)| dt = |X_1|, \\ U = \operatorname{sign}[\operatorname{sign}(X_1) \int_0^T V_2(t) \operatorname{sign}(V_1(t)) dt - X_2]. \end{cases}$$

В данном случае совместное решение уравнений (8) и первого уравнения (11) определяет единственное значение параметра T , а значение U определяется из второго уравнения (11).

Для двух управляющих параметров синтез оптимального управления определяется уравнениями

$$(12) \begin{cases} \int_0^{T_1} (|V_{21}(t)| + |V_{22}(t)|) dt = |X_2|, \\ U_1 = \text{sign}[\text{sign}(X_2) \int_0^{T_1} [V_{11}(t)\text{sign}(V_{21}(t)) + V_{12}(t)\text{sign}(V_{22}(t))] dt - X_1]; \end{cases}$$

$$(13) \begin{cases} \int_0^{T_2} (|V_{11}(t)| + |V_{12}(t)|) dt = |X_1|, \\ U_2 = \text{sign}[\text{sign}(X_1) \int_0^{T_2} [V_{21}(t)\text{sign}(V_{11}(t)) + V_{22}(t)\text{sign}(V_{12}(t))] dt - X_2]. \end{cases}$$

В данном случае вычисление единственных значений параметров T_1 , T_2 производится путем совместного решения уравнения (8) и первых уравнений систем (12), (13), после чего значения $U_1(X)$, $U_2(X)$ определяются из вторых уравнений систем (12), (13).

Очевидно, интегрирование в системах (8), (12), (13) следует производить в ускоренном модельном времени, что и позволяет определять значения U_1 и U_2 как функции текущих координат объекта.

На рисунке 1 представлена схема контроллера в среде MATLAB-SIMULINK без модели сопряженной системы.

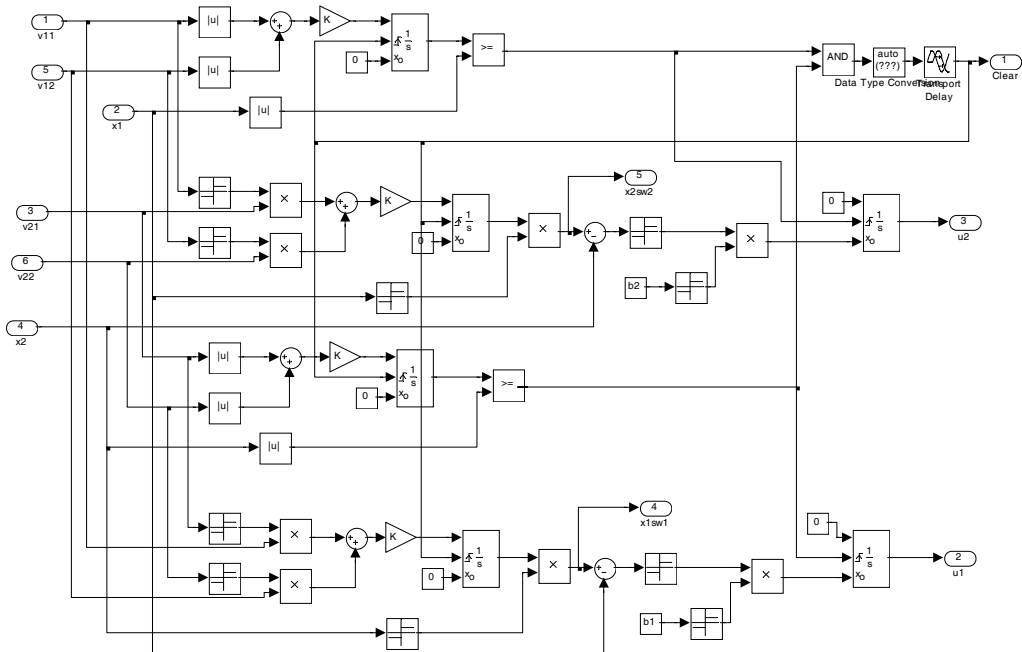


Рисунок 1 – Структурная схема контроллера.

На рисунке 2 приведены результаты моделирования оптимального управления для объекта с чисто мнимыми собственными значениями матрицы A . На этом же рисунке показаны линии переключения управляющих параметров.

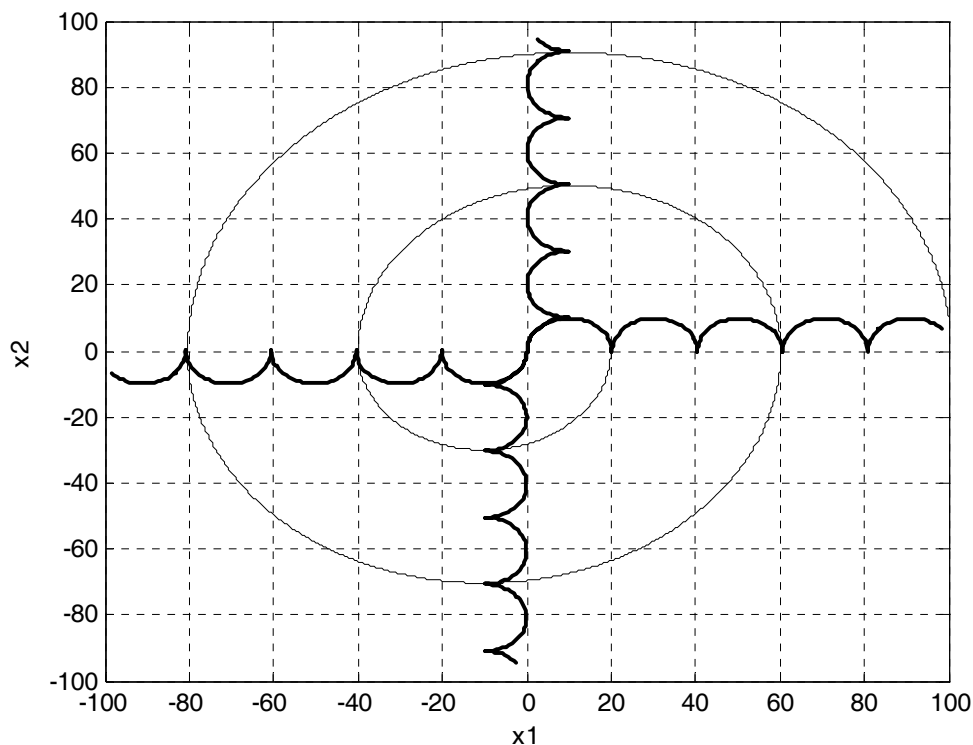


Рисунок 2 – Траектория движения управляемого объекта и линии переключения управляющих параметров.

Заключение

Предложен оптимальный по быстродействию регулятор, реализующий управление произвольным линейным объектом второго порядка в функции фазовых координат.

Список литературы

- [1] Понтрягин Л.С., Болтянский В.Г., Гамкредидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М., Наука, 1969.
- [2] Александров В.М., Нестеров А.А., Филиппова Н.П. Понижение размерности системы уравнений для синтеза оптимального управления. Оптимальные и самонастраивающиеся системы. – Новосибирск: Изд. ИАиЭ АН СССР, 1971. С. 17-26.
- [3] Нестеров А.А., Филиппова Н.П. Синтез оптимального управления для линейных объектов второго порядка. – Новосибирск: Изд. ИАиЭ АН СССР, 1971. с. 27-31.

УСТОЙЧИВОСТЬ ИНФОРМАЦИОННОГО ВЗАИМОДЕЙСТВИЯ В СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМАХ

Г.П. Аншаков, Я.А. Мостовой, А.В. Соллогуб

ФГУП «ГНИ РКЦ «ЦСКБ-Прогресс»

443009, г. Самара, ул. Псковская, 18

csdb@mail.samtel.ru

тел: (8462) 92-66-40, факс: (8462) 92-65-18

Ключевые слова: сложные технические системы, информационные взаимодействия, целостность информации, динамическая устойчивость, отказы и ошибки, информационная устойчивость

Abstract

Information processes in complex technical systems define a level of perfection and a direction of its development. Protection of the information is not reduced exclusively to protection against the non-authorized access to the information. Subjects of information attitudes can receive damage and from the termination of functioning of corresponding systems as a result of refusals and mistakes. Questions of a safety of work of the complex technical systems, arising are considered at consideration of consequences of the hardware and program mistakes interfering integrity of the information of system. These questions are examined from positions of information stability of system which is understood as property of system automatically to come back to normal functioning after passage of information indignations (distortions of information) from the specified reasons.

Введение

В настоящее время информационные процессы в сложных технических системах определяют уровень совершенства и поведения системы, направления её развития.

Безопасность информационных взаимодействий определяется защищенностью информации и поддерживающей инфраструктуры от случайных или преднамеренных воздействий естественного или искусственного характера, которые могут нанести неприемлемый ущерб субъектам информационных взаимодействий. Защита информации – комплекс мер, направленных на обеспечение безопасности информационных взаимодействий.

При этом термин «безопасность информационных взаимодействий» подразумевает наличие моделей угроз вычислительной системе, её ПО и данным, сети передачи данных и команд и соответствующих моделей защиты от этих угроз. Таким образом, информационная безопасность не может быть обеспечена повсеместно и в любых условиях, то есть имеются ограничения на область безопасности в рамках принятых моделей угроз и защиты от них.

Защита информации в сложных технических системах должна обеспечить:

- 1) Сохранение целостности информации – предотвращение несанкционированных изменений.
- 2) Сохранение конфиденциальности информации – предотвращение несанкционированного ознакомления.
- 3) Доступность информации – возможность для быстрого и санкционированного её изменения в оговоренное технологическое время.
- 4) Предотвращение несанкционированного использования (кражи) ресурсов системы.

Защита информации не сводится исключительно к защите от несанкционированного доступа к информации.

Субъекты информационных отношений могут получить ущерб и от прекращения функционирования соответствующих систем в результате отказов и ошибок, что в лучшем случае будет приводить к перерывам в работе системы.

Поэтому в рамках среды безопасности необходимо рассматривать пять составляющих: физическая безопасность, безопасность персонала, правовая безопасность, безопасность от информационных вторжений, безопасность функционирования, связанная с ошибками и отказами оборудования и ПО системы.

В настоящей статье рассмотрены вопросы обеспечения безопасности работы сложных технических систем, возникающие при рассмотрении последствий аппаратных и программных ошибок, приводящих к нарушению целостности информации системы. Эти вопросы рассматриваются с позиций информационной устойчивости системы, под которой мы будем понимать свойства системы автоматически возвращаться к нормальному функционированию после прекращения информационных возмущений (искажений информации) от указанных причин.

Данные задачи обеспечения информационной устойчивости решались при разработке систем управления космическими аппаратами. Однако аналогичные задачи и рассмотренные методы их решения возникают при разработке сложных систем более общего вида.

1 Защита информации в системе при сбоях и отказах аппаратуры. Информационная устойчивость

Методы обеспечения безопасности информационного взаимодействия в сложных технических системах (СТС) можно классифицировать по двум направлениям:

- 1) Недопущение информационных искажений – отражение информационных вторжений и попыток нарушения целостности информации.
- 2) Восстановление целостности информации коль скоро защита по п. 1 преодолена, и искажение информации произошло.

В первом случае необходимо обнаружение информационных атак, во втором случае – обнаружение информационных искажений. Защита информации от отказов и сбоев производится в рамках второго из указанных направлений.

Эксплуатация сложных систем показывает необходимость защиты информации системы от отказов и сбоев её структурных элементов, которые ведут к утрате целостности информации.

Обычно отказоустойчивость систем связывается с наличием либо аппаратной избыточности в ЦВМ, ПО, сетях передачи данных, либо наличием временной избыточности. Эта избыточность может быть использована несколькими различными способами:

- 1) путем допустимой деградации системы и её характеристик в пространстве работоспособных состояний при отказах и сбоях её структурных элементов;
- 2) восстановлением работоспособности отказавших структурных элементов путем использования взамен их резервных элементов;
- 3) повторными решениями одной и той же задачи по тому же или по другому алгоритму;
- 4) повторением передачи информации либо в качестве информационной обратной связи, либо прямое и n-кратное её повторение по схеме Вердана;
- 5) использование кодов обнаруживающих и устраняющих ошибки.

Однако наличие только резерва аппаратуры или временной избыточности не обеспечивает отказоустойчивости системы, также как вероятность безотказной работы (ВБР) не в полной мере её характеризует [3, 4].

Наряду с ВБР мерой отказоустойчивости должна быть вероятность сохранения целостности информации в системе, которая может нарушаться вследствие возникшего отказа или сбоя, а также вероятности парирования возмущений на систему, возникающих в момент восстановления её работоспособности.

Вопросы обеспечения информационной устойчивости составляют основное содержание понятия отказоустойчивости.

Таким образом, отказосбоеустойчивость сложных технических систем реализуется при наличии в системе следующих свойств [3]:

- 1) избыточности аппаратуры и в определённой мере ПО;
- 2) наличие средств встроенного контроля и диагностики для обнаружения и диагностики отказов и сбоев, а точнее - нарушения целостности информации;
- 3) наличия или сохранения «правильной» информации процессов управления для загрузки её в подключаемые резервные элементы при парировании сбоев или отказов аппаратуры системы.

При этом практическая реализация свойств 2) и 3) всегда сопровождается возмущениями на фазовые координаты и параметры системы, которые прикладываются к системе в момент возобновления правильного её функционирования. После восстановления работоспособности эти возмущения, связанные со старением информации в системе из-за отсутствия управления в течение некоторого времени, не должны приводить к потере устойчивости системой, то есть должны быть обработаны системой.

Величина возмущений фазовых координат и параметров системы зависит от принятого способа сохранения «правильной» информации в системе, а также от времени необходимого для распознавания нарушения целостности информации и времени извлечения «правильной» информации для её последующего использования.

Так как реальные СТС всегда нелинейны, то при достижении величины возмущений определенных значений возможна потеря устойчивости с переходом системы в новые устойчивые состояния.

Таким образом, анализ информационной устойчивости системы должен сопровождаться исследованием её динамической устойчивости.

Задача обеспечения устойчивости динамических систем многообразна и существует несколько определений устойчивости и условий её достижения: устойчивость к начальным возмущениям, устойчивость при постоянно действующих возмущениях, техническая устойчивость и т.п.

Учитывая это и то, что имеется связь между информационной и динамической устойчивостью в настоящей работе рассматривается задача устойчивости системы к начальным возмущениям и связанная с ней задача обеспечения защиты информации (информационной устойчивости) при отказах и сбоях аппаратуры системы.

Для получения условия информационной устойчивости рассмотрим два состояния сложной технической системы (рисунок 1).

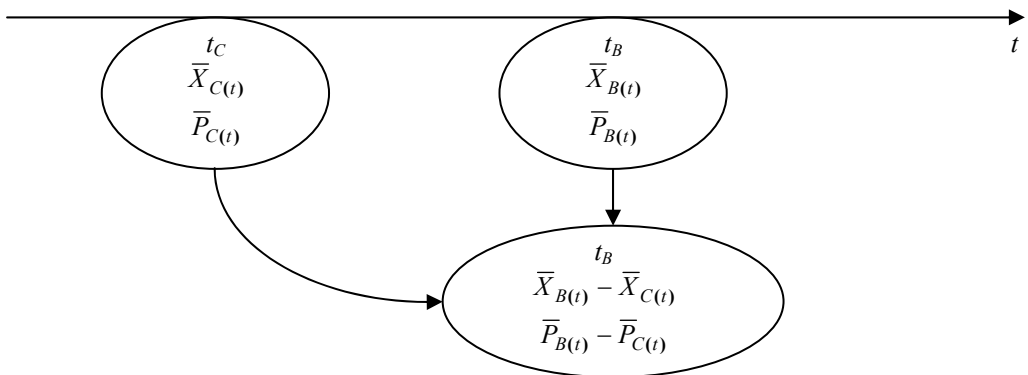


Рисунок 1 – Актуальные состояния сложной технической системы для анализа информационной устойчивости.

- 1) Состояние B – начало восстановления работоспособности системы после прошедшего отказа (сбоя) характеризуется временем t_B и совокупностью фазовых координат $\bar{X}_{B(t)}$ и параметров системы $\bar{P}_{B(t)}$.
- 2) Состояние C характеризуется ближайшим к t_B временем t_C и «правильной» совокупностью фазовых координат $\bar{X}_{C(t)}$ и параметров системы $\bar{P}_{C(t)}$.

При этом $t_B - t_C \geq \tau_{\text{обн}} + \tau_{\text{загр}}$, где:

- $\tau_{\text{обн}}$ – время, необходимое для обнаружения искажения информации;
- $\tau_{\text{загр}}$ – время, необходимое для извлечения и перезагрузки правильной информации.

Чем больше $\Delta t = t_B - t_C$, тем больше в общем случае $\Delta \bar{X}$, тем большие начальные возмущения действуют на систему в точке t_B . Некоторое уменьшение $\Delta \bar{X}$ возможно путем прогноза $\bar{X}_{C(t)}$ и $\bar{P}_{C(t)}$ на момент времени t_B .

Различные методы сохранения целостности - «правильности» информации в системе, обеспечивают соответственно различную вероятность сохранения устойчивости системы (различную вероятную величину возмущений на фазовые координаты системы) в момент восстановления управления.

2 Обнаружение искажения информации.

Методы автоматического восстановления правильной информации

Искажения информации наиболее просто обнаруживаются при её хранении или при передаче по линиям передачи данных. В устройствах, преобразующих информацию, например в процессоре ЦВМ, эти искажения обнаружить гораздо сложнее. Так как алгоритмы встроенного контроля и диагностики реализуются, а информация в системах преобразуется во встроенных ЦВМ, то вопросы информационной устойчивости сложных систем целесообразно рассматривать применительно к встроенным в сложные технические системы ЦВМ, как к критическому в смысле достижения информационной устойчивости звену.

Обнаружение отказов или сбоев, приводящих к искажению информации, может базироваться на обнаружении их последствий системными средствами - они проявятся через какой-то интервал времени в виде отклонений фазовых координат системы за допустимую при нормальном функционировании системы область. В результате команда на подключение резерва, загрузки в него «правильной» информации или на реконфигурацию будет получена с большой задержкой, когда продолжать работу нельзя и надо думать уже об аварийной защите системы, если она ещё возможна.

Поэтому целесообразно обнаружение искажений информации не связывать с ожиданием проявления ошибочных состояний системы. С этой целью часто применяется сравнение на аппаратном или программном уровне двух или более однородных результатов, полученных в управляющей вычислительной системе по одним и тем же исходным данным по одному и тому же алгоритму, либо по различным алгоритмам, но решающим одну и ту же задачу. Используются также аппаратный либо тестовый самоконтроль устройств системы или их взаимный контроль путем обмена контрольной информацией. Возможности современных вычислительных средств позволяют эффективно использовать моделирование будущих отдаленных состояний в качестве средств, определяющих искажение информации в текущий момент. В некоторых случаях эффективным оказывается контроль трафика системы. Данные методы обнаружения искажений информации обладают существенно меньшим запаздыванием, чем первый, то есть обнаружение нарушения целостности информации может произойти до того как они или их последствия стали различимы в поведении системы.

Где же взять «правильную» информацию в системе после обнаруженного её искажения?

При решении вопроса защиты информации процессов управления от последствий сбоя или отказа для продолжения работы системы после парирования отказа возможен вариант с «откатом» - возвратом процесса к точке, где целостность информации была обеспечена. Для возможности «отката» нужно запоминать и хранить информацию о состояниях процесса в потенциальных точках возврата - контрольных точках.

Также возможен вариант с «пропуском» - продолжением процессов управления после восстановления работоспособности по текущей информации в точке устранения неисправности или сбоя. Это справедливо для случаев процессов без памяти.

Возможен вариант «исправления» информации путем замены неверной информации в устройствах системы на правильную из заведомо работоспособных и параллельно работающих устройств. Возможно исправление информации из системы более высокого уровня иерархии.

Методы аппаратного мажорирования (голосования) при исполнении каждой машинной операции в троированной ЦВМ системы, а также избыточного отказоустойчивого кодирования при хранении и передаче информации обеспечивают использование избыточности, обнаружение ошибки, восстановление правильной информации в течение одной машинной операции на аппаратном уровне и не требует ни «отката» ни «пропуска».

Поэтому данные методы скорее относятся к методам предотвращения нарушения целостности информации, чем к методам её восстановления после искажения.

Тактика исполнения контрольной процедуры диагностики отказа (сбоя) должна учитывать последовательность выдачи и приема команд и информации во встроенной ЦВМ, обращений её к базе данных.

Без учета этого исполнение «отката» или «пропуска» не обеспечивает правильного функционирования системы в дальнейшем.

3 Аварийная защита

В общем случае «откат» с целью получения правильной информации для дальнейшей работы системы после сбоя или отказа можно осуществлять в другое устойчивое состояние системы. Эта возможность приводит к методу обеспечения работоспособности систем при отказах, который назовем «аварийной защитой».

Этот метод связан с организацией дополнительного устойчивого состояния системы, при нахождении в котором отказ может быть устранен и её работоспособность восстановлена. Дополнительные устойчивые состояния системы, возможные хотя бы для части наиболее значимых её фазовых координат, и методы перевода в них должны быть определены при проектировании системы. Например, для КА дополнительному устойчивому состоянию соответствует движение по орбите ИСЗ без угловой ориентации его связанных осей.

Рассматриваемая далее «аварийная защита» базируется на «мягком останове» системы при возникновении отказов и обнаружении искажений информации и переводе её в дополнительное устойчивое состояние «первоначального запуска».

«Мягкий останов» должен обеспечивать, как можно более организованное и приближенное к штатному выключение аппаратуры системы, что препятствует развитию аварийной ситуации и обеспечивает отсутствие необратимых последствий от отказа для системы, окружающей среды и информации. Этим самым «мягкий останов» создает условия для восстановления работоспособности системы и дальнейшего её устойчивого функционирования после проведения соответствующих ремонтных мероприятий.

Обнаружение нарушения целостности информации и её восстановление в сложной технической системе может занимать определенное время. Ещё большее время может занимать восстановление работоспособности и рестарт системы, например, путем внесения изменений в программное обеспечение. Использование аварийной защиты «с мягким останом» позволяет сочетать быструю реакцию системы на отказ с подключением к диагностике и восстановлению

работоспособности системы эксплуатирующего персонала или системы более высокого уровня иерархии.

Заключение

Безопасность информационных взаимодействий в сложных системах определяется защищенностью информации и поддерживающей инфраструктуры от случайных либо преднамеренных воздействий, которые могут нанести ущерб субъектам информационных воздействий.

Защита информации в сложных технических системах не сводится исключительно к защите от несанкционированного доступа к ней. Субъекты информационных отношений могут нести ущерб и от прекращения функционирования систем в результате отказов аппаратуры и ошибок ПО.

Рассматривается два направления защиты информации:

- недопущение нарушений целостности информации в системе путем обнаружения атак;
- восстановление искаженной информации в системе путем обнаружения искажений и придания системе свойств информационной устойчивости.

Метод обеспечения информационной устойчивости рассмотрен для случаев нарушения целостности информации системы при отказах (сбоях) аппаратуры и ошибках ПО. Информационная устойчивость в этом случае связана с исследованием динамической устойчивости системы.

Список литературы

- [1] Козлов В.Д., Аншаков Г.П., Мостовой Я.А., Соллогуб А.В. Управление космическими аппаратами зондирования Земли. Компьютерные технологии. – Москва: Машиностроение, 1998 г.- 386 с.
- [2] Антонов Ю.Г., Мостовой Я.А., Чайкин Ю.В. Принципы определения моментов проведения информационного согласования результатов работы каналов резервированной БВС КА / Тезисы докладов третьей международной научно-технической конференции «Микроэлектроника и информатика». - Москва, Зеленоград, 1997 г.
- [3] Anshakov G.P., Antonov Y.G., Mostovoy Y.A. Fault-tolerance and Disaster Protection of the Complex Technical Systems / International Symposium on Impact of Space Technology Innovation on Economic Development. - Shanghai, China, April 17-20 2001 г.
- [4] Аншаков Г.П., Антонов Ю.Г., Мостовой Я.А., Соллогуб А.В. Некоторые вопросы технологической защиты информации и информационной устойчивости сложных систем / Доклад на Научно-практической конференции «Проблемы информационной и экономической безопасности региона». – Самара: РАН, СГАУ. 2003 г.

ЭКСПЕРИМЕНТАЛЬНОЕ ПОДТВЕРЖДЕНИЕ МЕТОДА ОБНАРУЖЕНИЯ СРЫВНЫХ И ПОМПАЖНЫХ ЯВЛЕНИЙ ПО КОЛЕБАНИЯМ ЛОПАТОК

В.М. Белкин, В.П. Данильченко, Б.К. Райков, Ю.Н. Секисов, О.П. Скобелев

Институт проблем управления сложными системами РАН

443020, Самара, ул. Садовая, 61, Россия

sekisov@iccs.ru

тел: +7 (8462) 32-39-27, факс: +7 (8462) 33-27-70

Ключевые слова: *стендовые испытания, ввод двигателя в помпаж, метод обнаружения срывных и помпажных явлений, система измерения радиальных зазоров, мультистробоскопический опрос, критерии наличия срывных и помпажных явлений, количественные оценки критериев по результатам эксперимента, повышение чувствительности и быстродействия метода*

Abstract

The digital codes monitoring results at the output of system for radial clearances measuring in gas-turbine engines (GTE), provided by the method of stall and surge phenomena detecting on basis of blades vibration determination, are given. These results were obtained during GTE block tests when the engine worked at steady-state conditions, at transient conditions and during putting it in surge state. The estimated criterions based on experimental results are given. These criterions are the blade codes average values ratios during several periods of rotor revolution when the stall and surge phenomena are absent and exists. The way of sensitivity increase of proposed method is suggested. It efficiency is proved by experimental results. The measuring system use single-coil eddy-current sensors and their multistroboscopic scanning during several periods of GTE rotor revolution.

Введение

В работе [1] приводится описание метода обнаружения срывных и помпажных явлений в газотурбинном двигателе (ГТД) по колебаниям торцов лопаток в компрессоре. Метод предусматривает применение одновиткового вихретокового датчика (ОВТД) в составе системы измерения и контроля радиальных зазоров (РЗ) между торцами лопаток и торцом датчика, установленного на статоре, а также использование мультистробоскопического опроса датчика [2]. Предполагается мониторинг цифровых кодов в канале измерения зазора, а также определение текущих значений критерия, представляющего отношение средних значений кодов по лопаткам за несколько периодов вращения ротора в рабочем режиме работы двигателя и в состоянии помпажа.

Однако, описания метода имеют чисто теоретический характер, к тому же в них отсутствуют какие-либо экспериментальные подтверждения метода.

Поэтому в настоящей работе приводятся экспериментальные данные, полученные системой измерения РЗ при стендовых испытаниях компрессора ГТД как в рабочем режиме, так и в состоянии помпажа, а также количественные оценки критерия, найденные по результатам эксперимента. Кроме того, в работе предлагается способ повышения чувствительности метода обнаружения срывных и помпажных явлений и способ, обеспечивающий более высокое быстродействие в обнаружении этих явлений, эффективность которых также подтверждается экспериментальными данными.

1 Результаты эксперимента

В ходе стендовых испытаний компрессора высокого давления ГТД¹ изучались его характеристики в стационарных, переходных режимах и при помпаже. Наряду с традиционной измерительной аппаратурой в состав стендового оборудования была включена система измерения РЗ, использование которой предусматривалось программой испытаний. Подробное описание технических и программных средств системы измерения (СИ-01) приведено в работе [2], причем СИ-01 – это первая из серии систем измерения РЗ, программное обеспечение которой реализует мультитробоскопический опрос ОВТД. В соответствии с мультитробоскопическим опросом в течение каждого периода вращения ротора на измерительные цепи ОВТД подаются пакеты импульсов питающего напряжения, причем число импульсов в пакете равно числу контролируемых лопаток. На каждом последующем обороте ротора пакет импульсов смещается относительно предыдущего на шаг дискретизации. Число таких пакетов, равное числу оборотов ротора, в течение которых происходит преобразование, определяет его длительность. Выходные сигналы измерительных цепей преобразуются в цифровые коды, из которых формируются массивы данных по каждой лопатке. При отсутствии колебаний торцов лопаток в каждом массиве наблюдаются максимумы кодов, которые соответствуют искомому РЗ между торцами лопаток и чувствительным элементом (ЧЭ) датчика, установленного на статоре. По максимальным значениям кодов находятся физические значения РЗ, которые фиксируются в выходных протоколах системы.² Однако, в ходе испытаний компрессора в стационарном и квазистационарном режимах (которые соответствуют штатному режиму эксплуатации системы СИ-01) производился ввод двигателя в состояние помпажа для изучения поведения измерительных каналов СИ-01. При этом регистрировались не только максимумы кодов, по которым в соответствии с мультитробоскопическим методом вычислялись РЗ, но и все остальные результаты цифрового преобразования³. Эти результаты содержат важную информацию для подтверждения возможности практического использования метода обнаружения срывных и помпажных явлений по колебаниям лопаток, предложенного в работе [1].

ОВТД были установлены на 5 и 9 ступенях компрессора, причем результаты преобразования фиксировались по 5 ступени на каждой четвертой лопатке, а по 9 ступени на каждой пятой.

Ввод двигателя в помпажное состояние осуществлялся путем поджатия компрессора до срыва с помощью регулируемого сопла. Сначала устанавливался заданный режим по частоте вращения ротора с полностью открытым соплом, а затем с помощью стендовой системы управления уменьшалась площадь сопла вплоть до потери устойчивости.

При возникновении срывных и помпажных явлений торцы лопаток совершают колебательное движение и отклоняются от положений, характерных для бессрывного и беспомпажного состояния компрессора. При наличии строгой синхронизации опроса ОВТД с положением замков лопаток относительно датчика колеблющиеся торцы лопаток в моменты опроса датчиков оказываются в непредсказуемых положениях относительно своих замков, а результаты измерений – хаотичными [1].

На рисунке 1 представлены две группы зависимостей изменений кода как функции номера пакета (периода) для лопаток ($i = 1, 2, \dots, 21$) на 5 ступени. Первая группа ($C_{1i}(k)$) соответствует стационарному режиму (рисунок 1, а), вторая ($C_{2i}(k)$) отражает начало срывных и помпажных явлений (рисунок 1, б).

¹ Имеется в виду компрессор двигателя НК-86, серия испытаний которого проводилась на стендах НПО «Труд» (ныне СНТК им. Н.Д. Кузнецова) в 1988 г.

² Именно эти данные нашли отражение в отчетах предприятия, составленных по материалам испытаний двигателя НК-86.

³ Полученные данные длительное время оставались необработанными и в отчет предприятия по материалам испытаний двигателя НК-86 не попали. В то же время они сохранились в архиве разработчиков СИ-01 – сотрудников Самарского филиала Института машиноведения (ныне Институт проблем управления сложными системами РАН).

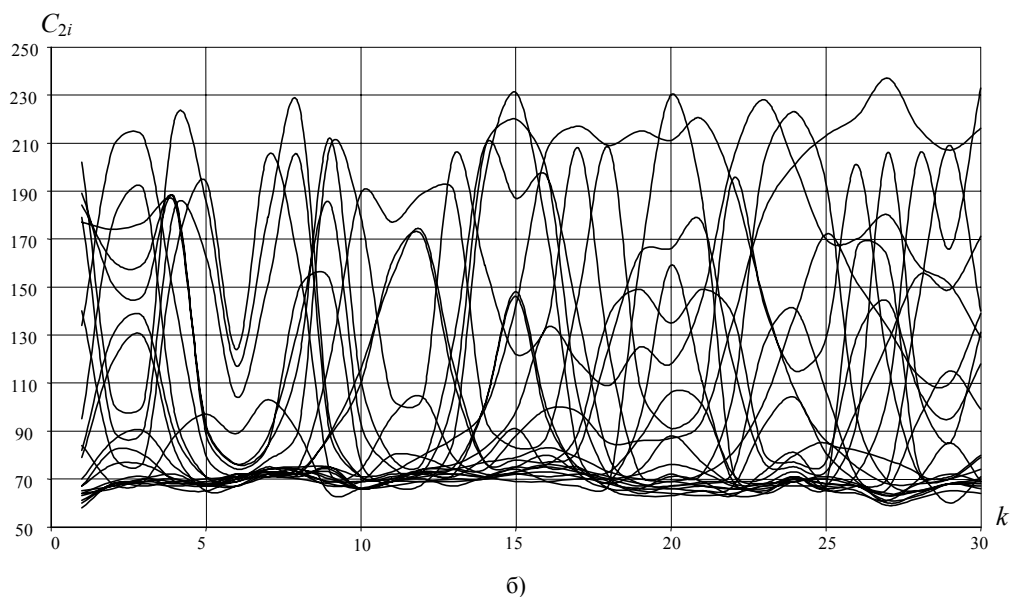
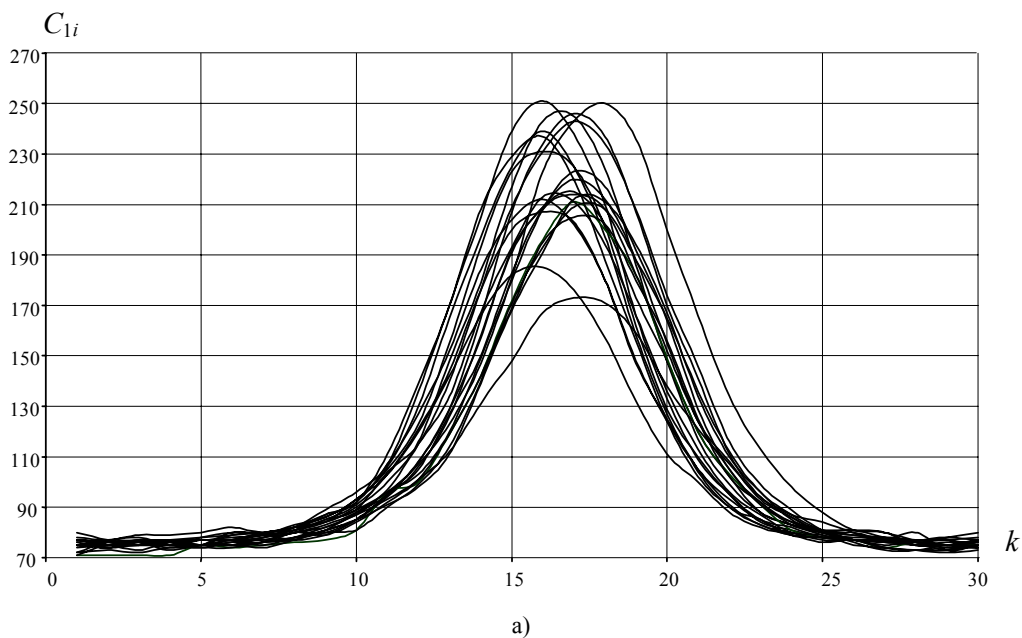


Рисунок 1 – Зависимости кодов от номера пакета (периода) при отсутствии ($C_{1i}(k)$) и наличии ($C_{2i}(k)$) срывных и помпажных явлений для группы контролируемых лопаток ($i=1, 2, \dots, 21$, ступень 5, замеры 31, 32).

Из графиков следует, что при стационарном режиме (отсутствии срывных и помпажных явлений) зависимости $C_1(k)$ для всех лопаток имеют идентичный характер с явно выраженным максимумом при сравнительно небольшом разбросе и отклонениях друг от друга. Напротив, появление колебаний лопаток, связанных со срывными и помпажными явлениями, нарушает прежнюю идентичность и приводит к хаотичному изменению кодов от номера пакета (периода) ($C_2(k)$).

Аналогичные зависимости получены для группы контролируемых лопаток ($i=1, 2, \dots, 21$) на 9 ступени, но в настоящей работе они не приводятся. Однако, полученные данные, как и данные, представленные на рисунке 1, использованы при построении зависимостей средних значений кодов по числу лопаток от номера пакета (периода) на тех же ступенях компрессора – графиков с на рисунке 2, а, соответствующих 5 ступени, а на рисунке 2, б – 9 ступени.

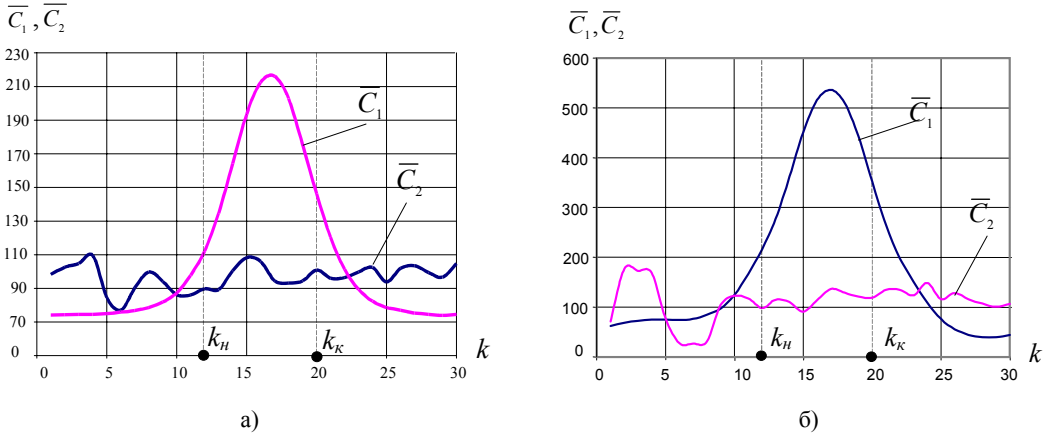


Рисунок 2 – Зависимости усредненных по лопаткам значений кодов от номера пакета (периода) при отсутствии ($\overline{C_1}(k)$) и наличии ($\overline{C_2}(k)$), срывных и помпажных явлений ($i=1, 2, \dots, 21$, ступень 5, замеры 31, 32 (а) и ступень 9, замеры 23, 24 (б))

Очевидно, что графики $\overline{C_1}(k)$ сохраняют форму исходных зависимостей $C_{1i}(k)$ и имеют явно выраженный максимум, который с появлением колебаний лопаток исчезает при сохранении незначительных пульсаций на графике $\overline{C_2}(k)$ (можно ожидать дальнейшее снижение пульсаций при увеличении числа контролируемых лопаток на обеих ступенях).

Вместе с тем, характер изменений функций $\overline{C_1}(k)$ и $\overline{C_2}(k)$, полученных по результатам эксперимента, в целом подтверждает теоретические положения метода обнаружения срывных и помпажных явлений, изложенного в работе [1]. Те же экспериментальные данные были использованы для оценки критерия

$$(1) \quad q = \frac{\sum_{i=1}^{n_i} \sum_{k=1}^K C_{1ik}}{\sum_{i=1}^{n_i} \sum_{k=1}^K C_{2ik}},$$

где C_{1ik}, C_{2ik} – коды при отсутствии и наличии колебаний. $\overline{C_1}(k)$

При $n_i=21$ и $K=30$ на 5 ступени критерий q составляет 1.15, а на 9 ступени – 1.7, причем согласно [1] пороговое значение критерия q ($q_{\text{пор}}$) выбирается между $q = 1$ и найденными числовыми значениями q .

В то же время различия в значениях q на обеих ступенях (при прочих равных условиях) весьма существенны и определяются как неидентичностью параметров измерительных каналов с ОВТД, так и различиями в поведении лопаток на обеих ступенях компрессора при срывных и помпажных явлениях. Поэтому окончательный выбор пороговых значений $q_{\text{пор}}$ должен учитывать индивидуальные особенности измерительных каналов и объекта.

Чувствительность метода к обнаружению срывных и помпажных явлений можно характеризовать разностью $q-1$, которая, судя по найденным числовым оценкам, невелика и на 5 ступени составляет всего 0.15.

Сравнительно низкая чувствительность и невысокое быстродействие (длительность процедуры составила 30 периодов вращения ротора ($K = 30$)) – это недостатки рассматриваемого метода.

2 Повышение чувствительности и быстродействия метода

Причиной снижения чувствительности метода к обнаружению срывных и помпажных явлений (колебаний лопаток) является нивелирование числителя и знаменателя выражения (1) с увеличением числа K , при котором критерий q уменьшается и стремится к единице.

В предлагаемом способе повышения чувствительности (п.ч.) число пакетов (периодов) ограничивается и выбирается равным разности $k_k - k_n$, где k_k, k_n – номера пакетов – конечного и начального соответственно (на рисунке 2 k_n и k_k ограничивают область наибольших значений \bar{C}). При этом выражение (1) принимает вид

$$(2) \quad q_{\text{п.ч.}} = \frac{\sum_{i=1}^{n_i} \sum_{k=k_n}^{k_k} C_{1ik}}{\sum_{i=1}^{n_i} \sum_{k=k_n}^{k_k} C_{2ik}}.$$

При сужении диапазона изменений k и выборе $k_n = 12$ и $k_k = 20$ значения $q_{\text{п.ч.}}$, найденные с помощью выражения (2), возрастают до 1.66 на 5 ступени и 3.17 на 9 ступени. Это означает более чем трехкратное повышение чувствительности ($q_{\text{п.ч.}} - 1$) на обеих ступенях компрессора.

Анализируя возможности повышения быстродействия (п.б.) метода, необходимо отметить, что обнаружение колебаний лопаток в принципе возможно и за один период вращения ротора при использовании только одного пакета. В этом случае фиксация кодов в канале ОВТД должна осуществляться в моменты прохождения центрами замков лопаток относительно ЧЭ датчика.

Если колебания отсутствуют, а изгибы пера лопаток, связанные с силовым воздействием газового потока, пренебрежимо малы, то значения кодов для всех лопаток будут максимальны.

Если колебаниями охвачены все лопатки, то вероятность попадания торца лопатки в зону чувствительности датчика существенно снижается, и значения кодов в те же моменты времени уменьшаются.

В качестве критерия начала колебаний можно использовать выражение

$$(3) \quad q_{\text{п.б.}} = \frac{\sum_{i=1}^{n_i} C_{1i(\text{з.л.})}}{\sum_{i=1}^{n_i} C_{2i(\text{з.л.})}},$$

где $C_{1i(\text{з.л.})}$ и $C_{2i(\text{з.л.})}$ – коды в момент прохождения центрами замков лопаток (з.л.) мимо ЧЭ датчика при отсутствии и наличии колебаний соответственно.

На рисунке 3 представлены отсчеты C_{1i} и C_{2i} , полученные в течение одного периода вращения ротора ($k = 17$) на 9 ступени в моменты прохождения центров замков лопаток $i = 1, 2, \dots, 21$ ЧЭ ОВТД. Аналогичные данные получены и на 5 ступени. И те и другие данные использованы для расчета критерия $q_{\text{п.б.}}$ (выражение (3)). Его значения достигают 2.28 на 5 ступени и 3.94 на 9 ступени.

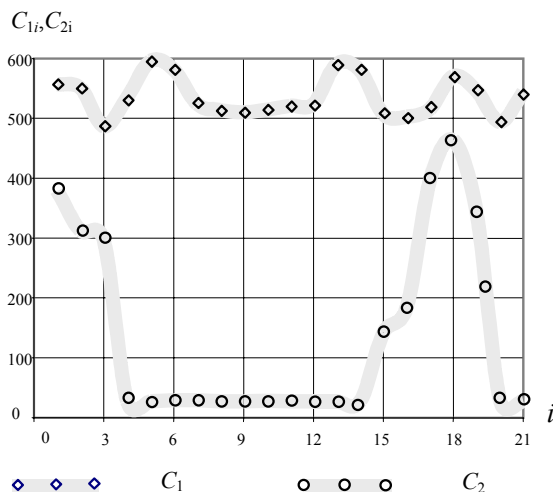


Рисунок 3 – Результаты отсчетов C_{1i} и C_{2i} за один период вращения ротора ($k=17$, ступень 9, замеры 23, 24)

Заключение

Результаты эксперимента, полученные в процессе стендовых испытаний компрессора ГТД, подтвердили основные теоретические положения метода обнаружения срывных и помпажных явлений по колебаниям лопаток и возможность его практической реализации. В то же время количественные оценки используемого в методе критерия, найденные по экспериментальным данным, показали сравнительно невысокую чувствительность метода к контролируемым явлениям при низком быстродействии. Предложены способы повышения чувствительности и быстродействия, которые также подтверждаются экспериментальными данными.

Список литературы

- [1] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Способы измерения и обнаружения колебаний лопаток в экспериментальных исследованиях и диагностике срывных и помпажных явлений в компрессорах газотурбинных двигателей // Проблемы управления и моделирования в сложных системах. Труды IV Международной конференции. – Самара: Самарский научный центр РАН, 2002, с. 539-545.
- [2] Методы и средства измерения многомерных перемещений элементов конструкции силовых установок / Под ред. Ю.Н. Секисова, О.П. Скобелева. – Самара: Самарский научный центр РАН, 2001 – 188 с.

ОДНОВИТКОВЫЕ ВИХРЕТОКОВЫЕ ДАТЧИКИ: ОТ КЛАСТЕРНЫХ КОМПОЗИЦИЙ К КЛАСТЕРНЫМ КОНСТРУКЦИЯМ

Л.Б. Беленький, Б.К. Райков, Ю.Н. Секисов, О.П. Скобелев

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
sekisov@iccs.ru

тел: +7 (8462) 33-31-31, 33-26-77, факс: +7 (8462) 33-27-70

***Ключевые слова:** одновитковые вихретоковые датчики, многокоординатные смещения элементов конструкций силовых установок, кластерные композиции датчиков, кластерные конструкции датчиков*

Abstract

Structures of existing eddy-current single-coil sensors with a sensitive element made as a piece of a conductor are systematized. Different combinations of sensors sensitive elements allocation inside cluster compositions meant for measuring coordinate components of multidimensional displacements of gas turbine and propeller fan engines blades are considered. Problems in constructing of cluster compositions by means of existent sensors are analyzed. Structures named as cluster single-coil eddy-current sensors are proposed. The description of the one of such cluster structure is given.

Введение

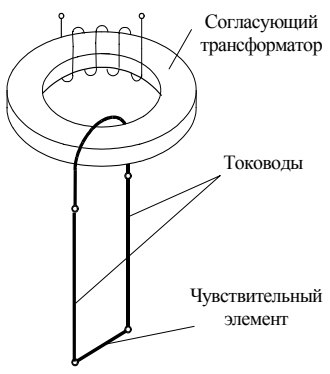
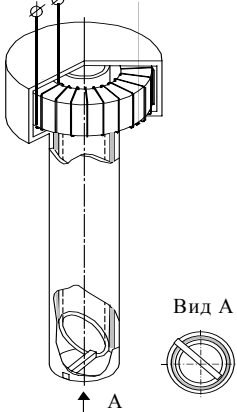
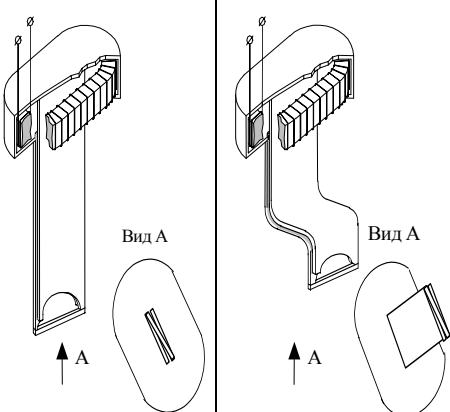
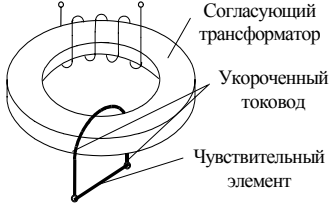
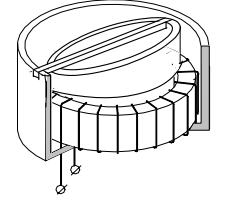
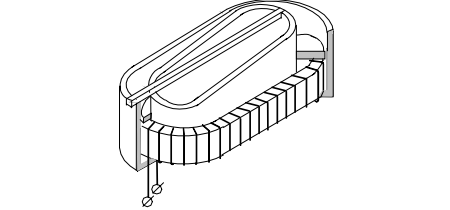
Первые публикации, связанные с одновитковыми вихретоковыми датчиками (ОВТД), чувствительные элементы (ЧЭ) которых имеют простейшую геометрическую форму в виде линейного отрезка проводника, обеспечивающего работоспособность датчика и повторяемость его характеристик при воздействии дестабилизирующих факторов, появились более 10 лет назад. Различные варианты конструкций ОВТД и измерительных цепей, включающих ОВТД, рассматриваются в работах [1, 2], а результаты их теоретических исследований приведены в работах [3-5]. В ряде публикаций [2, 5-8] приведены описания методов и средств измерения многокоординатных смещений элементов конструкций лопаточных и поршневых силовых установок, которые основаны на использовании ОВТД с ЧЭ в виде отрезка проводника.

Разнообразие существующих конструкций ОВТД вызывает необходимость их систематизации и обобщения, а также анализа трудностей и ограничений в построении кластерных композиций ОВТД, предназначенных для измерения многокоординатных смещений элементов конструкций силовых установок, причем проведение такого анализа способствовало бы разработке нового подхода к конструированию ОВТД и созданию перспективных датчиков. Однако, в известных литературных источниках [1-8] перечисленным вопросам уделено недостаточное внимание и поэтому в настоящей работе предпринята попытка устранения существующего пробела. В работе систематизируются известные конструкции ОВТД с ЧЭ в виде отрезка проводника, а на основе анализа существующих трудностей и ограничений в создании кластерных композиций ОВТД, предназначенных для измерения перемещений торцов лопаток турбомашин и лопастей винтовентиляторов, предлагаются перспективные конструкции, названные кластерными ОВТД.

1 Систематизация существующих конструктивных разновидностей ОВТД

Принцип действия ОВТД иллюстрируют схематические изображения, представленные в таблице 1 (колонка 1). ОВТД состоит из трех элементов: согласующего трансформатора, безындуктивных тоководов и ЧЭ. С помощью удлиненных безындуктивных тоководов ЧЭ вносятся в зону измерения с неблагоприятными внешними условиями, в частности в зону с высокой температурой (до 1200⁰С [1, 2, 9]). Согласующий трансформатор расположен в приемлемых условиях, а его первичная обмотка включается в измерительную цепь.

Таблица 1 - Одновитковые вихревые датчики с чувствительными элементами в виде отрезка проводника.

Принцип действия	Варианты механической конфигурации	
	Токовод цилиндрической формы	Плоские тоководы
1	2	3
		
		

Ток, протекающий в ЧЭ, создает магнитный поток, который, проникая в электропроводный объект, возбуждает вихревые токи, уменьшающие магнитный поток и индуктивность ЧЭ, а, следовательно, эквивалентную индуктивность датчика. Влияние вихревых токов зависит от расстояния до поверхности объекта, которое является входным преобразуемым параметром ОВТД.

Если ЧЭ датчика работает в условиях, приемлемых для магнитопровода, то необходимость в удлиненном токовом отпадает. В датчиках с укороченным тоководом (вторая строка таблицы 1) согласующий трансформатор находится вблизи зоны измерения [1, 2].

В таблице 1 (колонки 2, 3) представлены варианты механической конфигурации таких датчиков.

В наиболее распространенном варианте (колонка 2) удлиненные тоководы выполнены в виде изолированных друг от друга коаксиальных цилиндров, которые вместе с корпусом датчика образуют виток вторичной обмотки согласующего трансформатора. ЧЭ датчика замыкает оба токовода.

В варианте с укороченными тоководами возможен предельный случай, когда один из тоководов практически отсутствует, а ЧЭ непосредственно соединен с корпусом датчика (изображение этого варианта по сравнению с первым повернуто на 180 град. и ЧЭ находится в верхнем положении).

Следует отметить, что для решения многих практических задач важную роль играет степень локализации электромагнитного поля ЧЭ. В рассмотренных вариантах она зависит от экранирующего действия тоководов и корпуса датчиков и определяется их диаметрами.

В конструкциях ОВТД, представленных в колонке 3, удлиненные тоководы выполнены в виде близко расположенных и изолированных друг от друга жестких плоских или гибких проводящих полос [5, 10]. Гибкость тоководов при вводе ЧЭ в зону измерений обеспечивает преодоление (обход) возможных конструктивных препятствий на объекте.

Использование в ОВТД сердечников согласующих трансформаторов эллипсоидной формы [5] уменьшает их габариты и способствует сближению ЧЭ в составе кластерных композиций ОВТД.

Завершая описание механических конфигураций существующих ОВТД, необходимо подчеркнуть, что для преобразования сигналов таких датчиков используются измерительные цепи с импульсным питанием, в которых реализован известный метод первой производной [2], причем наибольшее распространение получили мосты Блумлейна и LR-мосты с дифференцирующими усилителями. Обычно в мосты включаются рабочий и компенсационный ОВТД для уменьшения влияния мешающих факторов, например, температуры.

2 Кластерные композиции ОВТД для измерения многокоординатных смещений элементов конструкций силовых установок

Известно, что экономичность и надежность современных лопаточных и поршневых силовых установок в значительной степени зависит от радиальных зазоров (РЗ) между торцами лопаток и статором, поршнем и цилиндром, от зазоров в подшипниках скольжения. В процессе функционирования элементы конструкции совершают многокоординатные смещения относительно друг друга, а РЗ является лишь одной из координатных составляющих (КС) таких смещений. Существующие методы измерения КС элементов конструкций в лопаточных и поршневых силовых установках предусматривают использование различных кластерных композиций ОВТД, в которых количество и расположение ЧЭ зависит от характера измеряемых смещений и определяется числом искоемых КС в выбранной системе отсчета. Описания кластерных методов измерения радиальных, осевых и угловых смещений торцов лопаток в компрессорах и турбинах приведены в работах [2, 6], методов измерения радиальных, осевых и изгибных, а также радиальных, осевых, изгибных и угловых смещений торцов лопаток и лопастей винтовентилятора (ВВ) авиационного газотурбинного двигателя (ГТД) – в работах [7, 8], причем в составе кластерной композиции используется от двух до четырех ОВТД (в зависимости от метода).

В работах [2, 5] рассматриваются метод измерения смещений оси поршня в трехмерном пространстве (шесть ОВТД) и метод измерения смещений центра вкладыша шатунного подшипника (ШП) с помощью множества ОВТД. Поршень и вкладыш ШП – это элементы конструкции кривошипно-шатунного механизма (КШМ) двигателя внутреннего сгорания.

Перечисленные методы обеспечивают наиболее полное представление о многокоординатных смещениях элементов конструкций, в том числе и о смещениях в радиальном направлении (РЗ). Для их реализации требуется, как уже отмечалось, определенное расположение ЧЭ датчиков в зоне измерения.

Рисунок 1 иллюстрирует одну из композиций ОВТД, которая предназначена для трехкоординатного (X, Y, Z) измерения смещений торцов лопаток в процессе их колебаний, сопровождающих срывные и помпажные явления в компрессоре ГТД. В соответствии с методом измерения [7] определяются КС смещений материальной точки M_i в момент прохождения замком лопатки геометрического центра, расположенного на равном расстоянии от ЧЭ₁ и ЧЭ₃. На рисунке 1 представлена также система отсчета ($0, X, Y, Z$) и рабочие диапазоны координатных смещений (X_p, Y_p, Z_p).

Если осуществляются только контрольные операции по обнаружению колебаний лопаток, то согласно [11] достаточно двух ОВТД в составе композиции (например, с ЧЭ₁ и ЧЭ₃) и даже одного датчика. Но в тех случаях, когда используются кластерные композиции ОВТД, методы предусматривают совокупную обработку результатов преобразования сигналов датчиков для нахождения искомым КС¹.

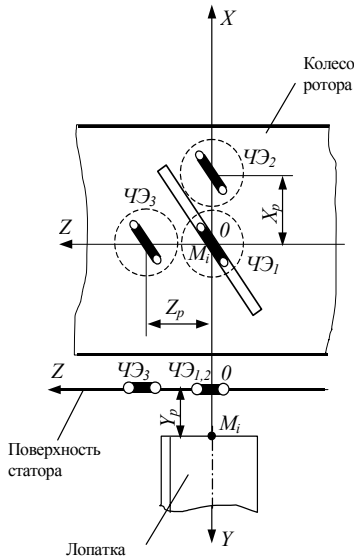


Рисунок 1 – Пример расположения ЧЭ в составе композиции ОВТД.

В системах измерения, обслуживающих КШМ, ЧЭ датчиков, как правило, расположены на координатных осях и КС определяются непосредственно с помощью градуировочных характеристик измерительных каналов системы по результатам преобразования в каждом канале [2, 5].

Однако, использование кластерных композиций связано с определенными трудностями и ограничениями, которые вызваны, главным образом, необходимостью выполнения установочных отверстий для ввода ЧЭ датчиков в зону измерений, а также конструктивными особенностями существующих ОВТД.

Установочные отверстия обычно выполняются (сверлятся) на конструктивно завершеном изделии в процессе его подготовки к испытаниям (исследованиям). Число отверстий возрастает с увеличением числа измеряемых КС и удваивается при необходимости использования дополнительных ОВТД для компенсации влияния мешающих факторов (в частности, температуры). Отверстия снижают надежность конструкции, и разработчики силовых установок стремятся минимизировать их количество.

Следует отметить, что при проведении измерений КС смещений торцов лопаток (лопастей) и использовании в составе композиции наиболее апробированных конструкций ОВТД их удлиненные тоководы, имеющие цилиндрическую форму (таблица 1), не позволяют разместить центры ЧЭ датчиков на расстоянии меньшем, чем диаметр тоководов. Более того, это расстояние дополнительно возрастает за счет согласующих трансформаторов, объемного витка, образующего верхнюю часть корпуса датчика (пунктирные окружности на рисунке 1 обозначают габаритные границы корпусов датчиков). Кроме того, тоководы оказывают экранирующее действие для электромагнитных полей ЧЭ датчиков и необходимы специальные доработки существующих конструкций для использования в составе кластерных композиций².

¹ КС смещений торцов лопаток определяются в результате решения системы уравнений в виде градуировочных характеристик измерительных каналов системы, найденных экспериментально и по результатам отсчетов в каждом канале [7].

² Подобные ограничения, хотя и возможны, но не типичны для кластеров ОВТД, используемых для исследований элементов конструкции КШМ, поскольку датчики, расположенные на координатных осях, обычно рассредоточены в пространстве [2, 5].

Кроме того, с увеличением числа датчиков в кластерной композиции возрастает разнесенность рабочих и компенсационных ОВТД, что неизбежно нарушает требование идентичности условий окружающей среды в точках размещения ЧЭ датчиков. Компенсация мешающих факторов становится неполной, что вызывает появление дополнительных погрешностей.

Использование ОВТД с плоскими гибкими тоководами и согласующими трансформаторами с магнитопроводами эллипсоидной формы позволяет в значительной степени избавиться от перечисленных ограничений. Однако, ряд нерешенных технологических задач, прежде всего, связанных с изготовлением тоководов, сдерживает построение кластеров на основе этой разновидности ОВТД.

Резюмируя, можно утверждать, что существующий подход к решению задач компенсации мешающих факторов, а также к построению кластерных композиций датчиков базируется на использовании функционально завершенных конструкций, рассчитанных на индивидуальное применение и монтаж на объекте, что и является причиной указанных трудностей и ограничений.

Основой новых разновидностей датчиков является антитетический, т. е. противоположный существующему, подход к их конструированию, который предусматривает создание единой (функционально завершенной) конфигурации, объединяющей несколько ОВТД. Такие конфигурации, называются далее кластерными одновитковыми вихретоковыми датчиками (КОВТД) (в отличие от кластера ОВТД [2, 4-8]).

3 Кластерные ОВТД

Конструкция простейшего кластерного датчика, приведенная в качестве примера на рисунке 2, объединяет два ОВТД из возможного множества. Ее практическое применение может быть в двух вариантах. В первом – один из ЧЭ используется как рабочий, другой ЧЭ – как компенсационный. Во втором варианте оба ЧЭ являются рабочими и предназначены для измерения двух КС (Y и Z).

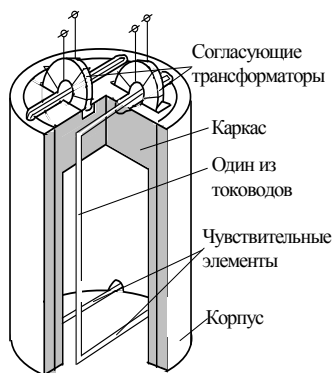


Рисунок 2 – Пример конструкции кластерного ОВТД.

Как показано на рисунке 2, собственно ЧЭ является нижняя часть витка, обращенная к торцам лопаток. Верхняя часть витка, проходящая через магнитопровод согласующего трансформатора, соединена с тоководами, которые покрыты слоем высокотемпературной изоляции и размещаются в пазах каркаса. Согласующий трансформатор и верхние части витков помещены в экраны. Экраны соединены с каркасом и корпусом, выполненным, как каркас и экран, из электропроводного немагнитного материала для уменьшения потоков рассеяния. Из-за экранирующего действия каркаса и корпуса индуктивность тоководов снижается по сравнению с индуктивностью ЧЭ (в идеале тоководы должны быть безындуктивными).

На рисунке 3 представлены положения торцов лопаток относительно ЧЭ₁, ЧЭ₂ КОВТД и координатных осей XYZ .

Если предположить, что в КОВТД в качестве рабочего используется ЧЭ₁, а компенсационным является ЧЭ₂, то при движении лопатки относительно датчика изменения индуктивностей L_1 и L_2 в зависимости от координаты Z будут иметь вид, изображенный на рисунке 4, а. При использовании датчика для измерения PZ (координаты Y) информативным параметром является минимум индуктивности, который соответствует моменту прохождения торца i -ой лопатки ЧЭ₁ (L_{i1min}). При этом индуктивность $L_2 = L_{20}$ и не зависит от смещения торца i -ой

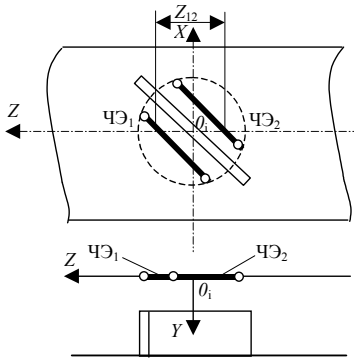


Рисунок 3 – Положение торцов лопаток относительно ЧЭ₁ и ЧЭ₂ и координатных осей XYZ.

лопатки по Y , а определяется состоянием окружающей среды. Желательно, чтобы расстояние между ЧЭ₁ и ЧЭ₂ было достаточно большим и изменения $L_1(Z)$ и $L_2(Z)$ в пределах $Z_{1,2}$ по возможности не «перекрывали» друг друга. Во втором варианте КОВТД, как отмечалось ранее, ЧЭ₁ и ЧЭ₂ являются рабочими и предназначены для измерения КС Y и Z . Расстояние между ЧЭ₁ и ЧЭ₂ уменьшается с таким расчетом, чтобы изменения $L_1(Z)$ и $L_2(Z)$ в пределах $Z_{1,2}$ были наибольшими (и «перекрывали» друг друга) (рисунок 4, б). Для компенсации влияния окружающей среды приходится использовать аналогичный датчик (с двумя ЧЭ), но, несмотря на это, число установочных отверстий сокращается вдвое.

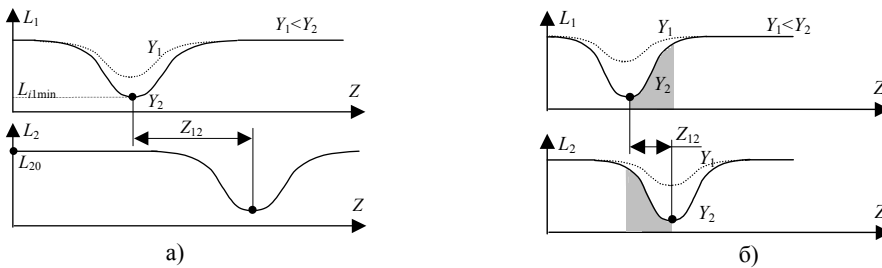


Рисунок 4 – Зависимость индуктивностей КОВТД от координаты Z : рабочий ЧЭ₁ и компенсационный ЧЭ₂ (а), рабочие ЧЭ₁ и ЧЭ₂ (б).

Заключение

Систематизированы и обобщены существующие конструктивные разновидности ОВТД. Проведен анализ трудностей и ограничений в создании кластерных композиций на основе существующих конструкций ОВТД, предназначенных для измерения многокоординатных смещений торцов лопаток турбомашин (лопастей винтовентиляторов).

В качестве основы перспективных разновидностей датчиков предлагается антитетический (противоположный) существующему подход к их конструированию, предусматривающий создание единой и функционально завершенной конструкции, объединяющей множество ОВТД, и названной кластерным ОВТД (КОВТД). Приводится описание простейшего примера конструкции КОВТД. Макетирование подтвердило работоспособность КОВТД, а проведенные исследования показали идентичность градуировочных характеристик с существующими разновидностями ОВТД [1, 2, 9] (при незначительном снижении чувствительности). Вместе с тем, требуется дополнительная отработка элементов конструкции, технологии изготовления. Необходимо также изготовление опытных образцов и испытания в широком диапазоне температур.

Список литературы

- [1] Райков Б.К., Секисов Ю.Н., Скобелев О.П., Хритин А.А. Вихретоковые датчики зазоров с чувствительными элементами в виде отрезка проводника // Приборы и системы управления, 1996, №8. С. 27-30.
- [2] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок / Под ред. Секисова Ю.Н., Скобелева О.П. – Самара: СНЦ РАН, 2001.
- [3] Хритин А.А. Система измерения радиальных зазоров в турбомашине: Дисс. ... канд. техн. наук. – Самара, 1993. – 174 с.
- [4] Секисов Ю.Н. Методы и средства измерений многомерных перемещений элементов конструкций силовых установок: Дис. ... докт. техн. наук. - Самара, 1999. - 267 с.
- [5] Беленький Л.Б. Системы измерения многомерных перемещений и деформаций элементов конструкций кривошипно-шатунного механизма в поршневых силовых установках. Дисс. канд. техн. наук. – Самара, 2003. - 198 с.
- [6] Боровик С.Ю., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Измерение и вычисление координатных составляющих многомерных перемещений торцов лопаток в процессе вращения ротора // Автометрия, 2001, №2. С. 103–111.
- [7] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Способ квазипараллельных измерений и вычислений координатных составляющих многомерных перемещений торцов лопаток // Проблемы управления и моделирования в сложных системах: Труды V Международной конференции – Самара: СНЦ РАН, 2003. С. 506–511.
- [8] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Метод измерения радиальных смещений лопастей винтовентилятора с использованием каналов физической и виртуальной коррекции // Проблемы управления и моделирования в сложных системах: Труды V Международной конференции – Самара: СНЦ РАН, 2003. С.512–520.
- [9] Патент № 1394912 МКИ G01N 27/90. Высокотемпературный проводниковый вихретоковый преобразователь/ Секисов Ю.Н., Скобелев О.П., Хритин А.А. - № 4136766/25-28; Заявл. 21.10.86; Оpubл. 03.05.95, Бюл. № 25.
- [10] Патент 2150676 РФ МКИ G01B 7/00. Вихретоковый преобразователь перемещений/ Беленький Л.Б., Секисов Ю.Н. - № 98108250/28; Заявл. 29.04.98; Оpubл. 10.06.2000, Бюл. №16.
- [11] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Способы измерения и обнаружения колебаний лопаток в экспериментальных исследованиях и диагностике срывных и помпажных явлений в компрессорах газотурбинных двигателей // Проблемы управления и моделирования в сложных системах: Труды IV Международной конференции – Самара: СНЦ РАН, 2002. С. 539–545.

СИСТЕМА ИЗМЕРЕНИЯ МНОГОМЕРНЫХ ПРОЦЕССОВ ДЛЯ ИССЛЕДОВАНИЙ И ДИАГНОСТИКИ СРЫВНЫХ И ПОМПАЖНЫХ ЯВЛЕНИЙ В ГАЗОТУРБИННОМ ДВИГАТЕЛЕ

С.Ю. Боровик, А.А. Инюцин, О.И. Камаева, В.В. Тулупова

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
borovik@iccs.ru, tulupova@iccs.ru
тел: +7 (8462) 33-26-77, факс: +7 (8462) 33-27-70

***Ключевые слова:** срывные и помпажные явления, многомерные процессы, система измерения, режим реального времени, кластерный метод измерения многомерных перемещений, обнаружение колебаний лопаток, диагностика срывных и помпажных явлений*

Abstract

The system for multivariate processes measuring in gas-turbine engine assigned for research and diagnostic of stall and surge phenomena is considered. The description of hard-and-software tools of the system as well as a detail description of it operational units and the results of their performance check are given.

Введение

В работе [1] предлагается метод диагностики срывных и помпажных явлений, отличительными особенностями которого являются бесконтактное измерение колебаний торцов лопаток в одном или нескольких сечениях компрессора с помощью одновитковых вихретоковых датчиков (ОВТД) и параллельные (одновременные) измерения параметров среды и режима объекта (оборотов ротора, тягового усилия, вибраций статора, расходов воздуха, давлений, температур). В работе [2] предлагаются способы обнаружения колебаний лопаток – по анализу изменений во времени координатных составляющих (КС) многомерных перемещений торцов лопаток, по смещению момента времени прохождения торцом чувствительного элемента датчика и с использованием статистической обработки измерительных кодов в зоне чувствительности ОВТД. Для измерения КС используется кластерный подход, основанный на определенном размещении в пространстве группы одинаковых датчиков (кластера) с заданной ориентацией их чувствительных элементов. Вычисление КС при этом осуществляется путем решения системы нелинейных уравнений, построенных на основе градуировочных характеристик измерительных каналов кластера [3]. Измерения, вычисления и диагностика должны осуществляться в режиме реального времени для каждой лопатки на каждом обороте ротора. Соблюдение этих требований возможно при использовании алгоритма квазипараллельных измерений и вычислений, в котором определение КС осуществляется в темпе получения кодов в каналах кластера [4].

В настоящей статье рассматривается разработанная авторами система измерения, реализующая основные положения предлагаемых методов и алгоритмов. Приводятся описания структуры технических средств системы и ее программного обеспечения, а также детализирующие описания отдельных функциональных блоков и результаты проверки их работоспособности.

1 Технические средства системы

В состав технических средств системы входят датчиковая аппаратура, платы L-Card (L-783 и E-440) для ввода аналоговых и цифровых сигналов, а также ПЭВМ (рис. 1). В обеих платах L-Card установлен сигнальный процессор с возможностью его программирования под прикладные задачи, что позволяет осуществлять распределенную обработку информации за счет выполнения ряда функций микропроцессорами.

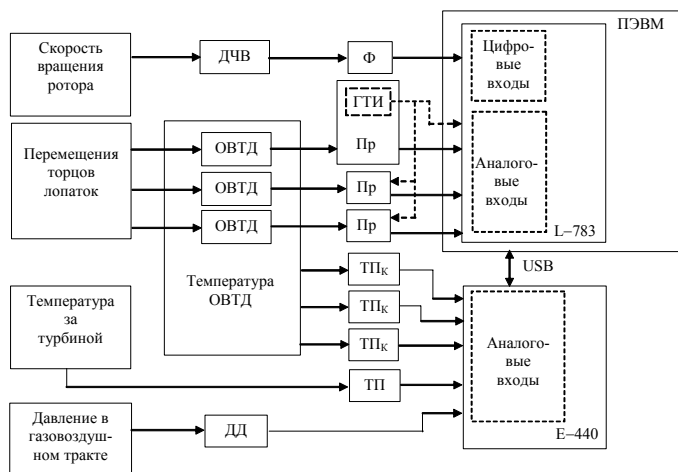


Рисунок 1 – Структура технических средств

назначены для обнаружения помпажных явлений традиционными методами по изменениям давлений на выходе ступени компрессора и по изменениям температуры на выходе турбины [1].

Вторая группа включает кластер ОВТД и преобразователи (Пр) естественных выходных сигналов ОВТД (индуктивностей) в нормализованные напряжения. Предусмотрена температурная коррекция сигналов ОВТД, которая производится с помощью термопар (ТПк), которые расположены в каждом датчике вблизи ЧЭ [3]¹. Сигналы ОВТД поступают на входы Пр [3], выполненных в виде отдельных блоков, соединенных с датчиками короткой линией (до 1 м). Генератор тактовых импульсов (ГТИ), входящий в состав одного из Пр, формирует управляющие сигналы импульсов питания ОВТД и импульсов внешней синхронизации АЦП платы L-783, обеспечивая тем самым синхронный запуск АЦП².

Выходные нормализованные сигналы с Пр подаются на аналоговые входы платы L-783, встроенной в ПЭВМ на шину PCI. На дискретный вход этой платы подается сформированный сигнал ДЧВ (для его преобразования используется таймер сигнального процессора). Сигналы оставшихся датчиков поступают на аналоговые входы внешней платы E-440, подключаемой к ПЭВМ по последовательному интерфейсу USB (платы E-440 и L-783 помимо АЦП содержат программируемый усилитель).

Предполагается, что система обслуживает одну ступень компрессора, на статоре которого установлен кластер ОВТД. Дополнительно измеряются обороты ротора, давление на входе или выходе ступени и температура газов на выходе турбины.

Датчиковая аппаратура содержит устройства промышленного изготовления и нестандартные датчики и преобразователи их сигналов. В состав первой группы входят датчик давления (ДД), температуры (термопары, ТП), датчик частоты вращения ротора (ДЧВ) и формирователь (Ф) его сигналов. ДД и ТП пред-

¹ Согласно методам, изложенным в работе [2], для обнаружения факта колебаний торцов лопаток (без измерения КС) достаточно и одного ОВТД.

² ГТИ имеется в составе каждого Пр и каждый из них можно использовать для синхронизации АЦП и ГТИ в остальных Пр.

2 Программное обеспечение системы

Программное обеспечение системы состоит из программ измерения КС, измерения параметров режима и среды объекта, обнаружения колебаний торцов лопаток, блока диагностики и программы визуализации.

Программа измерения КС функционирует в соответствии с алгоритмом квазипараллельных измерений и вычислений [4] и результатами ее работы являются значения координат по всем лопаткам на каждом обороте ротора. Предусмотрена также возможность функционирования в одноканальном варианте, реализующем диагностический режим работы системы (без измерения КС) [2]. Программа обнаружения колебаний выполняет проверку на факт наличия колебаний в соответствии с обоими вариантами программ. Одновременно осуществляется измерение периода вращения ротора, давления и температуры, а также приведение их к физическим значениям. Результаты измерений и вычислений анализируются в блоке диагностики, в котором по результатам анализа принимается решение о наличии срывных и помпажных явлениях. Программа визуализации работает в двух режимах – табличном и графическом. В первом режиме на экран выводится таблица значений КС по всем лопаткам, усредненных за 1 секунду, а также средние значения давления, температуры и скорости вращения ротора. Во втором режиме на экран выводятся графики изменения КС для любой выбранной лопатки с дискретностью в один период вращения ротора и графики изменения давления и температуры.

Все программы функционируют в реальном времени, дополнительно осуществляется непрерывная регистрация результатов в базу данных эксперимента.

Программный блок измерения КС и обнаружения колебаний торцов лопаток.

Реализация способа квазипараллельных измерений и вычислений [4] предъявляет высокие требования к быстродействию системы. На уровне сигнального процессора в течение периода тактовых импульсов внешней синхронизации (~1 мкс) осуществляется проверка текущего времени на соответствие моменту прохождения замком лопатки геометрического центра между чувствительными элементами датчиков, и при соответствии опрашиваются все три канала кластера. На уровне ПЭВМ в течение одного периода вращения ротора (5-20 мс) выполняется передача информации с плат АЦП, вычисляются и регистрируются КС по всем лопаткам, осуществляется контроль возможных колебаний торцов лопаток и диагностика срывных и помпажных явлений.

Структура программного блока представлена на рис. 2. В его состав входят управляющая программа сигнального процессора, драйвер для Windows 95/98/NT/2000 и программы прикладного уровня – вычисления КС и обнаружения колебаний.

Управляющая программа сигнального процессора работает по принципу команд. При поступлении команды из ПЭВМ командный монитор запускает ту или иную подпрограмму. С помощью команд задаются исходные данные для функционирования процесса сбора: массив логических номеров каналов, частота опроса АЦП, массив меток времени опроса каналов и ряд других параметров. С помощью логического номера канала задаются параметры сбора для каждого канала кластера: физический номер канала, тип подключения входных каскадов и коэффициент усиления. Второй массив используется для задания моментов времени опроса ОБТД, соответствующих прохождению торцом лопатки геометрического центра между датчиками [4]. При этом в качестве метки времени используется номер импульса внешней синхронизации, счет которого ведется от начала периода вращения ротора¹.

Подпрограмма опроса АЦП инициируется сигналом прерывания «Готовность АЦП». Она осуществляет управление адресом канала, коэффициентом усиления и отслеживает циклическую последовательность опроса, используя массив логических номеров каналов. Запуск АЦП для первого канала кадра осуществляется одновременно с приходом импульса синхронизации

¹ В качестве импульса внешней синхронизации используется сигнал ГТИ (рис. 1), длительностью порядка 1 мкс.

на соответствующий цифровой вход платы, если порядковый номер этого импульса равен текущей метке времени опроса из одноименного массива. Преобразование по следующему из списка каналу осуществляется через фиксированный интервал, равный запрограммированному периоду опроса АЦП. Преобразования повторяются для всех остальных каналов ОБТД в течение периода кадра, который равен периоду следования импульса синхронизации. Следует также отметить, что опрос каналов измерения КС осуществляется на максимально возможной частоте 3 МГц.

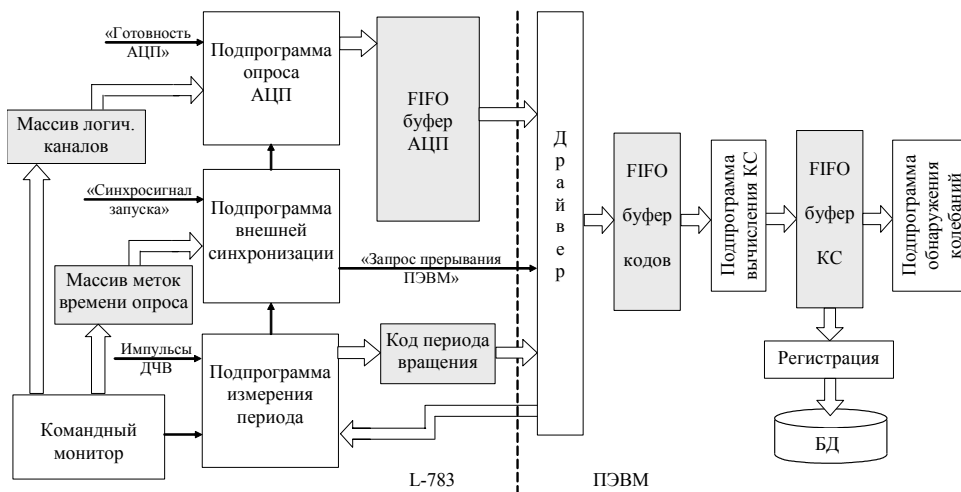


Рисунок 2 – Структура программного блока измерения КС и обнаружения колебаний торцов лопаток

Подпрограмма внешней синхронизации запускается с приходом каждого импульса синхронизации. Она выполняет счет этих импульсов и, если текущее значение счетчика совпадает с очередным значением метки времени, разрешает работу подпрограммы опроса АЦП. Сброс счетчика импульсов осуществляет подпрограмма измерения периода вращения, которая иницируется сигналом прерывания «Импульс ДЧВ» и измеряет временной интервал между прерываниями с помощью встроенного таймера. В качестве эталонной частоты, заполняющей измеряемый период, используется частота внутреннего генератора DSP, поделенная на некоторый коэффициент. Измеренный код периода вращения фиксируется в выделенной ячейке памяти, из которой считывается при отработке соответствующей команды от драйвера ПЭВМ для расчета меток времени опроса каналов.

Таким образом, в кольцевом FIFO буфере, реализованном во внутренней памяти сигнального процессора, регистрируются по 3 кода каналов кластера на каждую лопатку. При заполнении буфера генерируется прерывание в ПЭВМ. Драйвер платы по этим прерываниям считывает данные и помещает их в большой кольцевой FIFO буфер, реализованный в оперативной памяти компьютера. Программа вычисления КС обрабатывает данные в этом буфере, используя специальные указатели драйвера. Определение КС осуществляется путем решения системы нелинейных уравнений по алгоритмам, приведенным в [3, 4]. Вычисленные значения помещаются в кольцевой FIFO буфер КС для контроля возможных колебаний лопатки и регистрируются в базе данных эксперимента, организованной в виде файловой системы на жестком диске компьютера. Программа обнаружения колебаний работает параллельно с остальными программами ПЭВМ и осуществляет выявление колебательного процесса по изменениям каждой КС – превышению заданных пороговых значений, числа периодов колебаний и числа лопаток, охваченных колебаниями, используя при этом данные из кольцевого буфера КС.

Программный блок принятия решения о срывных и помпажных явлениях. Очевидно, что для своевременной выдачи предупреждающих сигналов средства диагностики, входящие в состав системы, должны обладать высоким быстродействием. Учитывая это требование, различную физическую природу процессов, приводящих к срывным и помпажным явлениям, а также преимущественно экспертный характер правил принятия решения о наличии диагностируемого явления, наиболее приемлемым представляется использование современных методов искусственного интеллекта, и в частности, методов на основе нечеткого вывода.

В качестве инструментальной среды для реализации программного блока принятия решения использовался пакет разработки нечетких моделей и систем FuzzyTech 5.54k корпорации Inform Software [5]. Такой выбор был обусловлен открытым интерфейсом пакета, который обеспечивает связь с большинством других программ благодаря использованию механизмов DDE, DLL, и ActiveX, а также наличием собственного API-интерфейса (Application Programming Interface), позволяющего интегрировать пакет с другими программами, разрабатываемыми исходя из потребностей конечного пользователя.

На рис. 3 представлена структура программного блока принятия решения о предпомпажном и помпажном состояниях ГТД в нотации пакета FuzzyTech 5.54k.

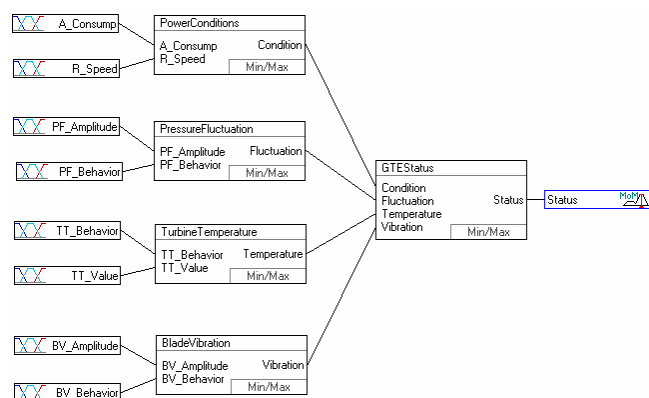


Рисунок 3 – Структура программного блока принятия решения о помпажном состоянии ГТД

«SurgePossibility» (возможность помпажа), «Foresurge» (предпомпажное состояние) и «Surge» (помпаж). Для дефазификации результата диагностики выбран метод поиска наиболее правдоподобного результата - метод центра максимума (CoM).

Входная переменная «Condition» формируется в блоке «PowerCondition» и определяет близость режима работы компрессора к линии границы помпажа в соответствии с его характеристикой. Нечеткий вывод осуществляется на основании информации о расходе воздуха через ступень компрессора и скорости вращения ротора, поступающей от измерительных каналов системы.

Входная переменная «Fluctuation» описывает характер пульсаций давления в компрессоре и определяется путем нечеткого вывода в блоке «PressureFluctuation» в соответствии с измененной амплитудой пульсаций и направлением ее изменения.

Аналогично, входная переменная «Temperature» блока «GTESStatus», описывающая характер изменений температуры за турбиной (разность текущей температуры от ожидаемой), определяется путем нечеткого вывода в блоке «TurbineTemperature». Исходной информацией, в этом случае, являются результаты измерения температуры за турбиной, сопоставленные с ожидаемой для заданного режима работы силовой установки, а также направление изменения указанной разности температур.

Заключение об угрозе помпажа формируется в блоке правил «GTESStatus» на основании контроля режима работы компрессора («Condition»), колебаний лопаток одного или нескольких рабочих колес («Vibration»), анализа давления на входе или выходе ступени компрессора («Fluctuation») и температуры за турбиной двигателя («Temperature»). Результатом оценки блока «GTESStatus» является лингвистическая переменная «Status», которая принимает значения «Stable» (стабильный режим работы установки), «Unstable» (нестабильный режим работы),

Наконец, переменная «Vibration», оценивающая колебания лопаток рабочего колеса диагностируемой ступени компрессора, определяется на основании блока правил «BladeVibration» в соответствии с измеренной амплитудой колебаний и направлением ее изменения (уменьшением или увеличением). При этом, предварительно определяется, соответствует ли частота колебаний лопаток диапазону «помпажных» частот (4-50 Гц).

Указанная на рис. 3 модель оценки интенсивности помпажных колебаний лопаток ориентирована, прежде всего, на способ измерения КС перемещений торцов лопаток и обнаружения их колебаний с помощью кластера ОВТД, позволяющего как выявить сам колебательный процесс, так и определить его характеристики. Реализации в системе способа обнаружения колебаний торцов лопаток с помощью одного ОВТД, обеспечивающего лишь фиксацию факта начала колебаний, вызывает необходимость ввода дополнительного термина входной лингвистической переменной «BV_Amplitude».

3 Результаты проверки работоспособности системы

В связи с жесткими требованиями по быстродействию, предъявляемыми к функционированию программного обеспечения в реальном времени, потребовалась экспериментальная проверка работоспособности разработанных программ в условиях, максимально приближенных к реальным. С этой целью использовались средства моделирования, описание которых приведено в работе [6]. Там же обосновывается необходимость их использования. На выходе моделей срывных и помпажных явлений в компрессоре, сопровождаемых колебаниями лопаток, формируются изменения параметров объекта во времени (давлений, температур, частоты вращения, а также КС перемещений торцов лопаток). Средства моделирования [6] обеспечивают также формирование напряжений на выходах ТП, ДД и ДЧВ, а также Пр, соответствующих изменениям перечисленных параметров. Полученные напряжения поступают на входы плат

L-783 и E-440 и по результатам их преобразования и обработки судят о работоспособности системы.¹

Экран результатов работы системы в графическом режиме приведен на рис. 4. В верхней части экрана показаны отсчеты КС (X, Y, Z) для одной из лопаток², а в нижней - отсчеты давления и температуры, измеренные на том же отрезке времени (пунктирные линии - гармонические колебания, получаемые с помощью модели). Изменения КС и пульсаций давлений носят колеба-

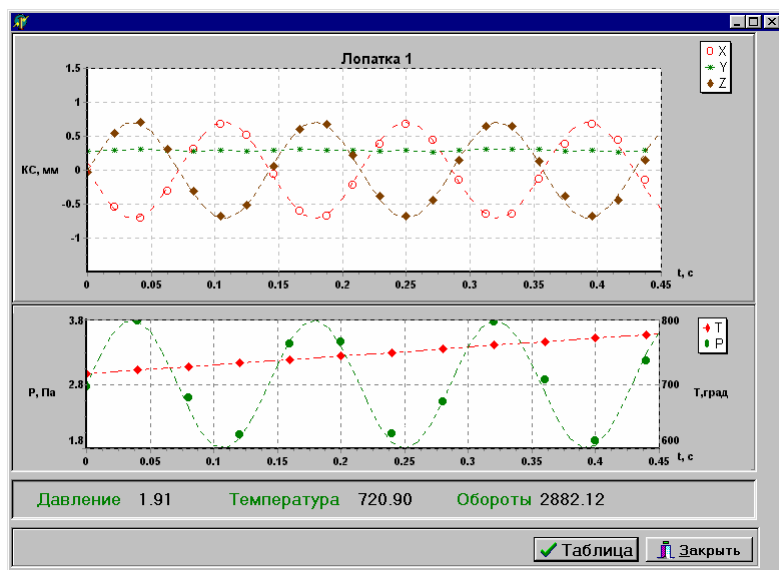


Рисунок 4 – Экран работы системы в графическом режиме просмотра

¹ Работоспособность датчиков и преобразователей промышленного изготовления не требует подтверждения, а наиболее полные сведения о работоспособности ОВТД и преобразователей их сигналов Пр имеются в работе [3].

² Аналогичные графики могут быть построены и для других лопаток (определение КС параллельно осуществляется для всех лопаток).

тельный характер, соответствующий результатам моделирования. При этом результаты, полученные в системе измерения и результаты моделирования практически совпадают.

Следует также отметить, что была проведена метрологическая аттестация каналов ОВТД (тип ОВТД-К-3(НС) [3]). Систематическая составляющая основной погрешности при аппроксимации градуировочной характеристики полиномом четвертой степени составила 0.5%, случайная составляющая менее 0.1%.

Заключение

Разработанная система позволяет осуществлять измерения КС перемещений торцов лопаток и обнаружение их колебаний в компрессоре, измерения давления и температуры в газоздушном тракте двигателя, а также проводить диагностику срывных и помпажных явлений. Технические решения, алгоритмы и реализующие их программы могут быть успешно применены при разработке подобных систем с большим числом каналов и большим набором выполняемых задач.

Экспериментальная проверка ранее не исследованных функций – измерения КС перемещений торцов лопаток и обнаружения их колебаний, подтвердила работоспособность системы.

Список литературы

- [1] Боровик С.Ю., Секисов Ю.Н., Скобелев О.П. Концептуальная модель измерения многомерных процессов в экспериментальных исследованиях и диагностике газотурбинных двигателей // Проблемы управления и моделирования в сложных системах. Труды IV Международной конференции. – Самара: Самарский научный центр РАН, 2002, с.531-538.
- [2] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П. Способы измерения и обнаружения колебаний лопаток в экспериментальных исследованиях и диагностике срывных и помпажных явлений в компрессорах газотурбинных двигателей // Проблемы управления и моделирования в сложных системах. Труды IV Международной конференции. – Самара: Самарский научный центр РАН, 2002, с.539-545.
- [3] Методы и средства измерения многомерных перемещений элементов конструкций силовых установок. / Под ред. Секисова Ю.Н., Скобелева О.П. – Самара: Самарский научный центр РАН, 2001. – 188 с.
- [4] Боровик С.Ю., Райков Б.К., Секисов Ю.Н., Скобелев О.П., Тулупова В.В. Способ квазипараллельных измерений и вычислений координатных составляющих многомерных перемещений торцов лопаток // Проблемы управления и моделирования в сложных системах. Труды V Международной конференции. – Самара: Самарский научный центр РАН, 2003, с.506-511.
- [5] А. Леоненков Нечеткое моделирование в среде MATLAB и fuzzyTECH. С-Пб.: БХВ - Санкт-Петербург, 2003. - 716 с.
- [6] Боровик С.Ю., Камаева О.И., Тулупова В.В. Средства моделирования для отладки и проверки работоспособности системы измерения многомерных проц// Проблемы управления и моделирования в сложных системах. Труды VI Международной конференции. – Самара: Самарский научный центр РАН, 2004 (настоящий сборник).

СРЕДСТВА МОДЕЛИРОВАНИЯ ДЛЯ ОТЛАДКИ И ПРОВЕРКИ РАБОТОСПОСОБНОСТИ СИСТЕМЫ ИЗМЕРЕНИЯ МНОГОМЕРНЫХ ПРОЦЕССОВ В ГАЗОТУРБИННОМ ДВИГАТЕЛЕ

С.Ю. Боровик, О.И. Камаева, В.В. Тулупова

Институт проблем управления сложными системами РАН
443020, Самара, ул. Садовая, 61, Россия
borovik@iccs.ru
тел: +7 (8462) 33-26-77, факс: +7 (8462) 33-27-70

Ключевые слова: срывные и помпажные явления, компрессор, газотурбинный двигатель, аппаратно-программные средства моделирования, система измерения, колебания лопаток, расход, давление, неустойчивая работа компрессора

Abstract

Modeling tools assigned for performance check and debugging of multivariate process measuring system in gas-turbine engine are considered. The descriptions of processes models that take place in power installation and in measuring system as well as the descriptions of soft-and-hardware tools, illustrated by model control screens, are given.

Введение

В работе [1] рассматривается система измерения многомерных процессов в газотурбинном двигателе (ГТД), которая содержит каналы измерения перемещений торцов рабочих лопаток колеса ротора компрессора ГТД и обеспечивает обнаружение их колебаний. В системе также предусмотрены каналы измерения параметров режима и среды объекта – скорости вращения, пульсаций давлений на выходе ступени и температуры газов на выходе турбины, что обеспечивает диагностику срывных и помпажных явлений в газоздушном тракте компрессора.

Одной из наиболее существенных трудностей при создании таких систем является отладка алгоритмов функционирования и реализующего их программного обеспечения, а также проверка работоспособности системы в целом. Использование натурального эксперимента, в процессе которого ГТД вводится в опасные состояния для отладки системы и проверки ее работоспособности практически нереально, в первую очередь по экономическим соображениям.

В этой связи, единственный путь для наладки и проверки работоспособности рассматриваемой системы - имитация эксперимента на моделях. В настоящей статье приводятся описания разработанных авторами аппаратно-программных средств моделирования, предназначенных для отладки и проверки работоспособности системы измерения многомерных процессов для экспериментальных исследований и диагностики срывных и помпажных явлений в ГТД.

1 Структурная схема средств моделирования

Реализация в системе измерения многомерных процессов [1] методов обнаружения колебаний лопаток за один период вращения ротора, предъявляет высокие требования к быстродействию как аппаратных, так и программных средств системы. Очевидно, что используемые для отладки и проверки работоспособности системы средства моделирования должны соответствовать этим требованиям. Однако, создание таких средств с единой программой управления расчетами и выдачей высокоскоростных сигналов затруднительно и вызывает необходимость использования дорогостоящей высокопроизводительной вычислительной техники.

Выходом из положения может служить подход, основанный на разнесении во времени процедур моделирования и выдачи сигналов, имитирующих работу датчиков и преобразователей. Другими словами, моделирование осуществляется в «мягком» времени при отсутствии каких-либо ограничений на скорость вычислений, а затем сформированные массивы данных выдаются на устройства вывода сигналов с высокой скоростью. При этом расчет модели и управление выдачей сигналов может осуществляться разными программными модулями.

Получение сигналов, имитирующих работу датчиков и преобразователей целесообразно осуществлять с использованием стандартных средств ввода/вывода, например плат фирм L-Card, National Instruments, Advantech и др. В то же время, для получения высокоскоростных сигналов необходима разработка специальных устройств.

На рис. 1 представлена обобщенная структурная схема аппаратно-программных средств моделирования срывных и помпажных явлений в ГТД, реализующих предложенный подход.

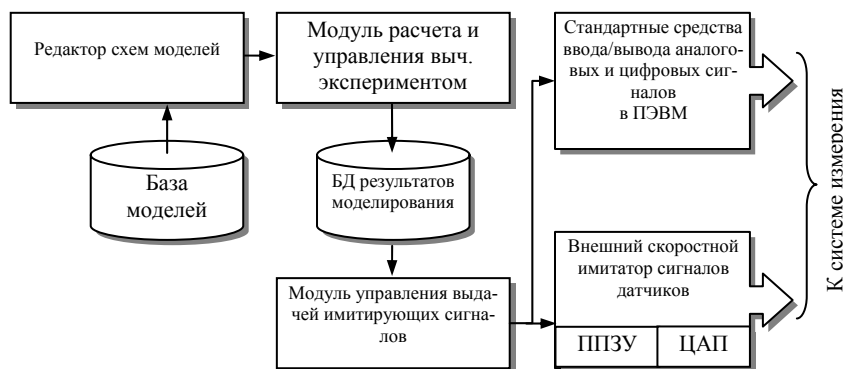


Рисунок 1 – Структура аппаратно-программных средств моделирования

В состав технических средств входит универсальная плата ввода/вывода аналоговых и цифровых сигналов L-783 (разработка фирмы L-Card), имеющая в своем составе 2-х каналный ЦАП и 16 цифровых выходов, а также специально разработанный внешний модуль, содержащий 3 банка ППЗУ и 3 одноканальных высокоскоростных цифро-аналоговых преобразователя (ЦАП).

Программное обеспечение включает редактор схем моделей, обеспечивающий формирование вычислительных схем на основе имеющихся в базе моделей; программный модуль расчета моделей и управления вычислительным экспериментом, обеспечивающий расчет сформированной редактором вычислительной схемы и формирование базы данных (БД) с результатами моделирования, а также программный модуль управления выдачей сигналов, осуществляющий считывание информации из БД результатов моделирования и выдачу ее на средства вывода имитирующих сигналов.

2 Модели процессов

В системе измерения [1] заключение об угрозе помпажа формируется на основании контроля колебаний рабочих лопаток одного или нескольких колес диагностируемой ступени компрессора, анализа пульсаций давления на выходе ступени и отклонений температуры за турбиной двигателя от нормы. Этот набор измеряемых параметров и определяет количественный и качественный состав моделей, которые должны быть включены в средства моделирования срывных и помпажных явлений в компрессоре ГТД.

Моделирование пульсаций давлений осуществляется на основе экспериментально определяемых семейств характеристик компрессора [2], устанавливающих связь между приведенным

весовым расходом воздуха, степенью повышения давления и приведенными оборотами ротора. Геометрическое место точек указанных минимальных расходов образует линии срыва или границы помпажа. В качестве входных переменных используются значения оборотов и расхода воздуха через ступень компрессора, а в качестве выходной переменной – давление на выходе ступени.

При расчете модели для заданных значений оборотов определяется соответствующая им характеристика зависимости давления от расхода воздуха и ищется точка пересечения данной характеристики с линией границы помпажа ($G_{помп.}$). Когда обороты ниже номинальных, определяется и точка пересечения характеристики с линией границы начала срыва ($G_{срыв.}$). Если заданный расход воздуха превышает $G_{помп.}$ и $G_{срыв.}$, то по найденной характеристике определяется давление на выходе ступени компрессора, которое будет постоянно во времени при заданных расходе и оборотах.

Если же обороты ниже номинальных и заданный расход воздуха находится в диапазоне $G_{срыв...}G_{помп.}$, то по найденной характеристике определяется средний уровень давления на выходе ступени компрессора и формируются высокочастотные гармонические изменения давления и расхода, характеризующие срыв. Для случая, когда заданный расход воздуха меньше $G_{помп.}$ - по найденной характеристике определяется средний уровень давления на выходе ступени компрессора, а на выходе модели формируются низкочастотные колебания давления, характеризующие помпаж. С целью упрощения моделируется только «мягкий помпаж» вблизи точки $G_{помп.}$, не имеющий разрывной характеристики [2]. Очевидно, что при оборотах ниже номинальных, произойдет наложение колебаний давления, характеризующих помпаж и вращающийся срыв.

Модель вращения лопаточного венца колеса ротора и возникновения колебаний лопаток обеспечивает формирование на выходе значений координатных составляющих (КС) перемещений торцов контролируемых лопаток в системе отсчета XYZ , жестко связанной со статором.

Если рассматривать вращательное движение лопаточного венца ротора турбокомпрессора, то перемещения контролируемого торца лопатки будет происходить только в радиальном направлении (вдоль оси Y) и в направлении вращения ротора (ось Z). Смещения же в направлении оси ротора (ось X) останутся неизменными. Перемещения торца каждой лопатки вдоль оси Z (проекция перемещения на ось Z) могут быть описаны простейшими тригонометрическими функциями. Перемещения торцов контролируемых лопаток вдоль оси Y определяются начальным значением радиального зазора и поворотом колеса ротора.

Аналогичным образом могут быть определены КС перемещений торцов лопаток при их колебаниях. С целью упрощения предполагается, что колебания лопаток имеют гармонический характер. При этом модель колебаний лопаток должна интегрироваться с моделью вращения колеса ротора.

Моделирование выходных сигналов датчиков и преобразователей параметров газовой турбины, а также параметров режима осуществляется с помощью набора моделей для каждого параметра. Исходными данными для моделирования являются градуировочные характеристики (ГХ), представленные в том или ином виде (например, коэффициенты полинома или интерполяционная таблица. Результатами моделирования процесса преобразования являются изменения напряжения (или кода), соответствующие изменениям входных параметров). Соответственно, исходными данными для *моделирования сигналов датчиков и преобразователей радиальных, осевых и изгибных смещений торцов лопаток* являются семейства ГХ для каждого ОВТД, координаты X , Y и Z , изменяющиеся во времени, полученные для каждой лопатки при наличии или отсутствии колебаний. Результаты моделирования представляют собой изменения напряжений или кодов во времени, соответствующие входным сигналам каждого ОВТД и преобразователя кластера в течение нескольких периодов вращения ротора компрессора ГТД.

3 Программное обеспечение формирования схем и расчета моделей

Используемые модели образуют базу моделей (рис. 1) и программно могут быть реализованы как в виде отдельных исполняемых модулей, так и в виде программ стандартных средств моделирования (например, Mathcad, MatLab и др.). Объединение моделей в единую вычислительную сеть осуществляется с помощью специального графического редактора схем моделей.

В качестве примера, на рис.2 представлена схема моделей для расчета пульсаций давлений (модель GTEStabOp), смещений торца лопатки при ее колебаниях с учетом вращения колеса ротора компрессора ГТД (модель BladeOsc), а также моделирования выходных сигналов преобразователей радиальных, осевых и изгибных смещений торцов лопаток (модель MeasSys). Все модели, включенные в схему, реализованы в виде отдельных программ.

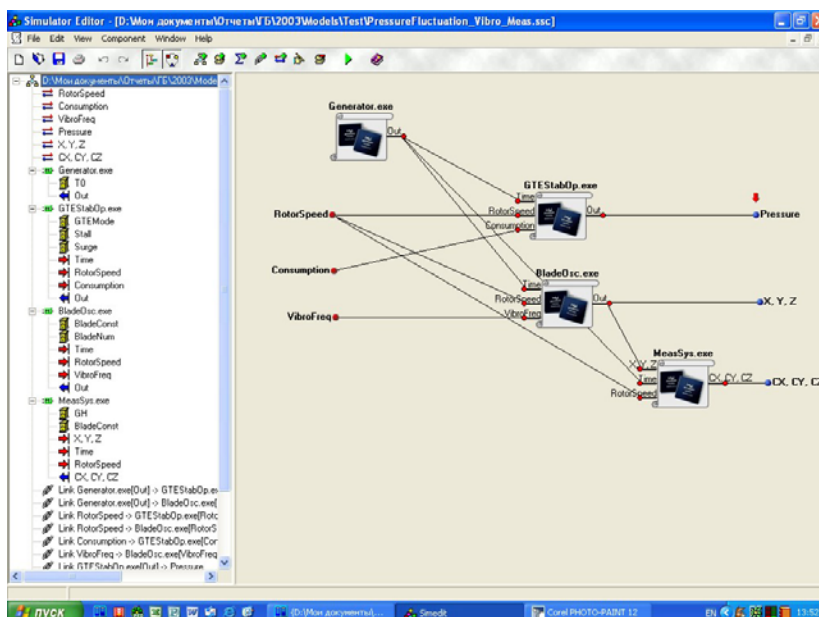


Рисунок 2 – Экран редактора схем моделей

Входами схемы моделей являются значения скорости вращения ротора (RotorSpeed), значение расхода воздуха через моделируемую ступень компрессора (Consumption) и частота колебаний выбранной лопатки колеса ротора ступени компрессора ГТД. Выходы схемы – пульсации давлений на выходе ступени (Pressure), изменения КС (X, Y, Z) выбранной лопатки, а также соответствующие им изменения кодов (CX, CY, CZ). Задание временных меток, для которых осуществляется расчет схемы, осуществляется с помощью специальной модели-генератора (Generator).

В разработанном программном обеспечении средств моделирования редактор схем моделей и программа расчета и управления вычислительным экспериментом (рис. 1) реализованы в виде отдельных модулей. Связь между ними осуществляется посредством файлов с описаниями последовательностей расчета моделей схемы. Сформированный редактором схем файл поступает на вход программы расчета и управления вычислительным экспериментом, где сначала производится настройка вычислительного эксперимента с помощью «электронного помощника», позволяющего задать способ ввода исходных данных (ручной или посредством файлов данных) и вид представления результатов моделирования (табличный, графический, комбинированный), а затем запускается собственно блок расчета схемы моделей.

На рис. 3 представлен экран работы программы расчета и управления вычислительным экспериментом в режиме расчета схемы моделей. На экране показаны графики пульсаций давления на выходе ступени компрессора ГТД в режиме помпажа (нижний график), а также изменения положения торца одной из лопаток колеса ротора вдоль оси ротора (верхний график) и в направлении его вращения (средний график) при колебаниях лопатки, вызванных помпажем. Здесь же представлена и специальная консоль ручного режима управления, установлением положений виртуальных регуляторов которой осуществляется изменение параметров расчета моделей.



Рисунок 3 – Экран блока расчета моделей

С целью исключения избыточной информации, расчет КС смещений торцов лопаток осуществляется только для моментов времени, соответствующих прохождению каждой лопатки в пространстве, ограниченном рабочим диапазоном ОВТД (кластера ОВТД). При выходе за указанную зону, КС смещений торцов лопаток присваиваются значения «по-умолчанию» ($X=Z=0$; $Y=Y_{\infty}$, где Y_{∞} - значение «бесконечного» зазора, соответствующее показаниям системы измерения при прохождении под чувствительным элементом ОВТД межлопаточного промежутка). Это и обуславливает «дискретный» характер графика смещений торца отдельно выбранной лопатки на рис. 3.

4 Программное и аппаратное обеспечение управления выдачей имитирующих сигналов

Как отмечалось в п. 2, имитация выходных сигналов датчиков давлений и температуры осуществляется с использованием универсальной платы ввода/вывода L-783. В составе платы имеется двухканальный 12-разрядный ЦАП с программируемой частотой преобразования до 125 кГц и внутренний буфер объемом до 2 кВ для хранения кадра выдаваемой через ЦАП информации. Управление микросхемой ЦАП и организацией кольцевого буфера кадра осуществляет установленный на плате сигнальный процессор.

Программа управления выдачей имитирующих сигналов пульсаций давлений и температуры на основе результатов моделирования формирует массив данных для ЦАП для каждого канала и загружает его в кольцевой буфер платы. Программа задает также необходимые параметры работы ЦАП (частоту преобразования и размер кадра), рассчитанные в соответствии с параметрами моделирования (шаг моделирования и период моделируемого сигнала). После того, как программа управления разрешила работу ЦАП, его функционирование осуществляется в автономном режиме.

В штатный BIOS (встроенное ПО платы) добавлена возможность синхронизации перезапуска выдачи информации через ЦАП с использованием внешнего цифрового синхроимпульса. Синхроимпульс заведен на вход внешнего прерывания сигнального процессора, при обслуживании запроса которого указатель текущего выдаваемого отсчета автоматически передвигается в начало кадра.

Имитация выходных сигналов датчика частоты вращения (ДЧВ) с формирователем и кластера ОВТД с преобразователями осуществляется в специально разработанном высокоскоростном внешнем модуле. Объединение генерации этих сигналов в одном модуле продиктовано необходимостью жесткой синхронизации выдачи сигналов, имитирующих работу кластера датчиков ОВТД и ДЧВ.

В состав модуля входят 3 высокоскоростных 12-разрядных ЦАП с частотой преобразования 1.7 МГц (по одному на каждый датчик кластера), 3 банка ППЗУ объемом 512 Кб, тактовый генератор и устройство управления. Устройство управления состоит из счетчика адреса памяти, формирователя запуска ЦАП и формирователя внешнего синхроимпульса для системы измерения. В начале каждого тактового импульса инкрементируется счетчик адреса, выдаваемые коды из каждого банка ППЗУ поступают параллельно на входные шины ЦАП, жестко закрепленными за этими ППЗУ и формируется с некоторым временным сдвигом (для завершения цифро-аналогового преобразования) синхроимпульс системы измерения.

Выходной сигнал, соответствующий частоте вращения, формируется с использованием двоичных данных, записанных в старшем (неиспользуемом для ЦАП) разряде одного из банков. Обновление данных моделирования осуществляется путем электрического перепрограммирования микросхем памяти.

Заключение

Рассмотренные аппаратно-программные средства моделирования срывных и помпажных явлений в компрессоре ГТД обеспечивают расчет измеряемых системой параметров, отражающих устойчивую работу компрессора и его функционирование в условиях срыва и помпажа. Объединение программ расчета моделей с аппаратными средствами имитации сигналов датчиков и преобразователей, используемых в системе, позволило осуществить отладку основных функциональных блоков системы, а также проверить работоспособность наиболее критичного по быстродействию блока измерения КС перемещений торцов лопаток и обнаружения их колебаний в лабораторных условиях при минимальных материальных затратах на проведение экспериментов.

Список литературы

- [1] Боровик С.Ю., Инюцин А.А., Камаева О.И., Тулупова В.В. Система измерения многомерных процессов для исследования и диагностики срывных и помпажных явлений в газотурбинном двигателе // Проблемы управления и моделирования в сложных системах. Труды VI Международной конференции. – Самара: Самарский научный центр РАН, 2004 (настоящий сборник).
- [2] Хьюперт М. Помпаж в компрессоре. Сб. переводов Аэродинамический расчет осевых компрессоров № 9185, том 2, институт им П.И. Баранова, М. 1964, 58 – 606 с.

АПРИОРНОЕ МОДЕЛИРОВАНИЕ В ЗАДАЧАХ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

В.И. Батищев

Самарский государственный технический университет
443100, Самара, ул. Молодогвардейская, 244, Россия
vib@list.ru
тел: +7 (8462) 37-12-74, факс: +7 (8462) 37-15-19

Ключевые слова: интерпретация, априорная информация, экспериментальные данные, измерения, контроль, диагностика

Abstract

Problems of using the approximate approach to measuring tasks are considered. The methods of parametric definition of apriori models and the problems of experimental data interpretation are analyzed.

Введение

Построение математических моделей по результатам измерений является конечной целью широкого круга исследовательских и диагностических задач. Измерения, испытания, контроль, идентификация, с одной стороны, различные по целям задачи, с другой стороны, все они связаны с получением информации об объекте экспериментальным путем. По сути своей они являются аппроксимационными задачами. Любая информация об объекте исследования, полученная на основе экспериментально определенных данных и представленная в виде чисел, графиков, формул есть модель исследуемой реальности, т. е. аппроксимирующая ее форма. Полнота, достоверность и полезность этой информации определяется как процедурой обработки экспериментальных данных, так и априорными сведениями об объекте исследования.

В последние десятилетия сформировалось направление, связанное с разработкой и построением измерительных систем, предполагающих включение математических моделей непосредственно в их рабочий контур [1, 2]. В рамках этого направления разработаны различные аппроксимационные ИИС для анализа вероятностно-статистических характеристик случайных процессов и полей, определения состава и свойств веществ, идентификации, контроля, диагностики и прогнозирования состояния технических систем [2-4]. При всем разнообразии рассмотренных выше задач важной и актуальной является проблема обобщения подходов к их решению. Использование аппроксимационного подхода, как общей платформы для построения методологического, математического и алгоритмического обеспечения ИИС, дает предпосылки системного объединения через математическую модель априорной информации, выражающей физическую сущность решаемой задачи, структуры и характеристик измерительных средств и методов, а также процедуры осуществления измерений и интерпретации их результатов. Основным достоинством таких систем является компактная, удобная для пользователя форма представления результатов, что позволяет их обоснованно интерпретировать и при необходимости включать непосредственно в процедуры дальнейших аналитических преобразований и расчетов.

Развитию теории и совершенствованию методологии прикладного аппроксимационного анализа способствуют работы, посвященные приближенно-аналитическим методам исследования. Эти методы используются для решения задач прикладной математики, математической физики, обработки данных и ряда других в тех случаях, когда получение точных решений не-

возможно или затруднено из-за недоопределенности и случайности исходных данных, некорректности постановки задач. Приближенно-аналитические методы лежат в основе разработки алгоритмического обеспечения измерительно-вычислительных и моделирующих комплексов и систем.

Следует отметить, что идея системного объединения измерительных и моделирующих ресурсов оказывается плодотворной по отношению к задачам испытаний, контроля и диагностики производственных процессов и объектов, транспорта в рабочих режимах и условиях нормального функционирования, а также изделий промышленного производства на протяжении всего их жизненного цикла.

1 Априорные модели в измерительных задачах

То, что называют аппроксимационным подходом, есть некоторое обобщение принципов, методов, способов и средств, направленных на построение явных аналитических моделей, вид которых выбирается на основе априорной информации и имеющегося фактического экспериментального материала с учетом целей проводимого исследования. Построение математических моделей динамических систем по результатам измерений составляет предмет теории идентификации. Сложность таких задач находится в прямой связи со сложностью объекта исследования, и при определенном ее уровне термин «идентификация» становится некорректным. Формирование моделей сложных систем связывается с теорией интерпретации результатов измерений и наблюдений.

Интенсивно развивающиеся методы и средства идентификации оказали существенное влияние на пересмотр основных положений теории измерений, информационно-измерительные технологии, и всю экспериментально-исследовательскую практику. Прежде всего это связано с признанием фундаментальности понятия «математическая модель» в теории измерений. «На фундаментальном уровне современная теория измерений занимается алгоритмизацией процесса измерения с учетом математического и физического аспектов понятия измеряемой величины... Моделированием ИИС обеспечивается фундаментальная строгость и обоснованность конкретных технических решений» [6].

Фундаментальным свойством аппроксимационного подхода является многовариантность порождаемых им методов и средств реализации, которая следует из многовариантности интерпретаций любых данных, в которых есть случайность и неопределенность.

Поскольку любая аппроксимация дает приближенный результат, в условиях многовариантности нет смысла искать наилучшие решения в отношении выбора моделей и методов их параметрического доопределения. Выбор вариантов решения задачи следует осуществлять по принципу непротиворечивости результата и его свойств целям и задачам исследования. Этим требованиям отвечают интерпретационные модели.

Задача выбора и обоснования вариантов при построении аппроксимационных ИИС осложняется еще и многокритериальностью такого выбора. При проведении измерения значения физической величины цели и критерии предельно ясны: получить как можно более точный результат как можно более простыми средствами за кратчайший отрезок времени. В рассматриваемом случае исходные предпосылки, методы и средства достижения результата должны быть согласованы и с исходными данными – случайными и недоопределенными, и с априорной информацией, и с достигаемыми целями. Естественно, критерии такого согласования могут быть различными, а зачастую – противоречивыми.

Особую роль априорная информация играет при решении обратных задач. В принципе, разнообразие методов решения обратных задач определяется характером и полнотой привлекаемых дополнительных сведений об искомом решении. В информационно-измерительной технике с решением обратных задач связаны процедуры коррекции результатов измерений, направленные на устранение влияния аппаратной функции (редукция к идеальному прибору) и

разделение эффектов от влияния отдельных факторов, а также процедуры интерпретации результатов косвенных измерений в спектроскопии, дифрактометрии, хроматографии, томографии.

Исследования в области теории и приложений обратных задач шли, в основном, по пути поиска общих, универсальных решений. Существенный прорыв в этом направлении произошел с появлением и интенсивным развитием методов регуляризации, базирующихся на привлечении априорной информации общего характера, например, неотрицательности и гладкости решений. Аппроксимационные методы требуют более полной и структурированной априорной информации, что делает возможным их применение лишь в узких конкретных прикладных задачах. Однако в таких ситуациях методы аппроксимации позволяют получать устойчивые, адекватные исходным данным и работоспособные решения значительно более простыми средствами, чем при использовании методов регуляризации. Привлечение априорных моделей позволяет делать принципиально новые постановки задач и получать прямые методы решения обратных задач [7].

Наконец, в пользу более пристального внимания к аппроксимационным методам решения обратных задач говорит то, что в различных областях, таких как аналитическая химия, рентгеновская дифрактометрия, γ -спектроскопия, гидродинамика, известно большое количество работ, посвященных обоснованию аналитических математических моделей, свойственных и адекватных различным объектам, процессам и ситуациям [8]. Наличие такой модельной базы и разработка соответствующих методов и алгоритмов оптимального параметрического доопределения моделей составляют основу для существенного расширения функциональных возможностей и областей применения информационно-измерительных и контрольно-диагностических средств.

Задачи промышленного контроля, испытаний и диагностики связаны с необходимостью проведения измерений при рабочих режимах функционирования оборудования, в связи с чем возникают проблемы, обусловленные случайным характером и зачастую ограниченным объемом экспериментальных данных, необходимостью получения результатов обработки и принятия решений в минимальные сроки. Решению таких проблем способствует также привлечение априорной информации в форме аналитических моделей исследуемых зависимостей и разработка соответствующих статистических методов обработки и анализа данных.

2 Методы параметрического доопределения априорных моделей

В ряде задач контроля и испытаний вид сигнала строго обусловлен физическими законами исследуемых явлений, а погрешности измерений пренебрежимо малы. Примерами могут служить переходные процессы и периодические сигналы в электрических цепях. По принципу анализа свойств переходных процессов работают системы испытаний и контроля блоков электронной аппаратуры, электро-энергетических сетей и оборудования. На основе измерения характеристик периодических сигналов осуществляется контроль разного рода электрических и электронных генераторов, оценивается качество электрической энергии, проводятся испытания колебательных механических систем, например, маятниковых часовых механизмов.

Известны способы определения характеристик таких сигналов, основанные на обработке результатов измерений некоторого количества их мгновенных значений. Все эти способы имеют обобщение в рамках методов аппроксимации, обеспечивающего равенство значений сигнала $x(t)$ и его модели $x_M(t, \alpha_1, \dots, \alpha_m)$ в выбранных точках t .

При различных, в общем случае, произвольных значениях аргумента t проводятся измерения m значений сигнала, и составляется система m уравнений:

$$(1) \quad x_M(t_i, \alpha_1, \dots, \alpha_m) = x(t_i), \quad i = 1, \dots, m,$$

которая решается относительно параметров $\alpha_1, \dots, \alpha_m$.

Поскольку в основу построения метода не заложено никакого метрологического критерия, основной проблемой на пути внедрения метода является анализ точности и обоснование достоверности получаемых результатов. При совпадении модели и сигнала получается методически точный результат. При выборе модели без должного обоснования и несоответствии ее виду моделируемого сигнала оценки параметров могут существенно отличаться от оптимальных в равномерном или среднеквадратическом смысле. Поэтому каждое конкретное приложение должно сопровождаться обстоятельным метрологическим анализом.

В общем случае качество решения аппроксимационной задачи оценивается некоторой результирующей погрешностью, включающей в себя все составляющие, определяющие несоответствие модели и моделируемой зависимости. При случайных сигналах это интегральные, как правило, среднеквадратические оценки. При детерминированных сигналах могут быть получены более жесткие оценки равномерного приближения.

Для получения аналитических оценок погрешностей предлагается использовать модели общего вида, включающие используемую интерпретационную модель как частный случай. Такими общими эталонными моделями могут служить функциональные ряды

$$x_{э\tau}(t) = \sum_{k=0}^N a_k \varphi_k(t), \quad N \rightarrow \infty.$$

Критерий среднеквадратического приближения функций и его модификации имеют ряд известных достоинств. Целесообразно использовать в качестве критерия приближения его обобщенный вариант – минимум взвешенной среднеквадратической погрешности аппроксимации анализируемой функции $R(\tau)$ ее моделью $R_M(\tau, \vec{\alpha})$ [1, 4]:

$$(2) \quad \min \left\{ \rho(R, R_M) = \int_a^b [R(\tau) - R_M(\tau, \vec{\alpha})]^2 \mu(\tau, \vec{\beta}) d\tau \right\},$$

где $\mu(\tau, \vec{\beta}) \geq 0$ – некоторая весовая функция с параметрами $\vec{\beta} = \beta_1, \dots$

Пределы интегрирования a и b в общем случае могут быть произвольными и зависят от свойств модели $R_M(\tau, \vec{\alpha})$, а также специфики решаемых задач.

Критерий (2), наряду с его очевидным достоинством, заключающимся в возможности выделения участков наиболее точного приближения, позволяет косвенно решить ряд частных задач анализа. Например, получая серию оценок параметров модели $R_M(\tau, \vec{\alpha})$ при различных значениях параметров $\vec{\beta}$ весовой функции, можно судить о принадлежности зависимости $R(\tau)$ к определенному классу.

При решении практических задач, в ряде случаев, конечной целью оценивания являются не параметры модели $R_M(\tau, \vec{\alpha})$, а некоторые характеристики, функционально связанные с параметрами $\vec{\alpha}$. В этом случае можно ставить задачу оптимального или рационального выбора вида весовой функции $\mu(\tau, \vec{\beta})$ и подстройки ее параметров $\vec{\beta}$.

Используя неадекватные упрощенные модели, при соответствующем выборе весовой функции $\mu(\tau, \vec{\beta})$ и ее параметров, можно определить характеристики локальных свойств исследуемой зависимости $R(\tau)$.

Другим, обобщающим большую группу применяемых на практике критериев, является критерий равенства значений функционалов: $P_i \{X(t), Y(t), \vec{\gamma}\}$ – экспериментально оцененных по

реализациям исследуемых процессов $X(t), Y(t)$ и $P_{Mi} \{R_M(\tau, \vec{\alpha}), \vec{\gamma}\}$ – аналитически вычисленных для выбранной модели $R_M(\tau, \vec{\alpha})$, т.е.

$$(3) P_i \{X(t), Y(t), \vec{\gamma}\} = P_{Mi} \{R_M(\tau, \vec{\alpha}), \vec{\gamma}\}, \quad i = 1, \dots, m,$$

где $\vec{\gamma} = \gamma_1, \dots, \gamma_N$ – параметры операторов $P_i\{\}$.

Оценки параметров $\alpha_1, \dots, \alpha_m$ определяются как решение системы уравнений (3). При использовании критерия (3) большие возможности исследователю дает выбор того или иного вида операторов $P_i\{\}$ и $P_{Mi}\{\}$, а также значений их параметров. Характерным частным случаем этого критерия является известный критерий моментов.

Заключение

Обоснованная интерпретация результатов экспериментальных исследований должна базироваться на априорной информации, формализованной, как правило, в виде априорной модели объекта исследования и апостериорной информации, полученной путем обработки результатов измерений. Априорная информация может быть, также, учтена при синтезе алгоритмов обработки экспериментальных данных, и оказать существенное влияние на их структуру и свойства.

Дальнейшее развитие математического моделирования и его приложений, на наш взгляд, пойдёт по пути автоматизации некоторых этапов создания математических моделей. Классификация и выбор моделей – это также процесс измерения, но в шкале наименований на основе измерения некоторой совокупности характеристик исследуемых объектов в признаковых шкалах. Такая методология построения систем призвана дополнить возможности компьютеров по эффективной работе с цифрами и символами возможностями эффективно работать с образами.

Список литературы

- [1] Батищев В.И. Измерительно-моделирующие методы оценивания функциональных характеристик случайных процессов: // Проблемы управления и моделирования в сложных системах: Труды IV международной конференции (Самара, 17-23 июня 2002 г.). – Самара: СНЦ РАН, 2002. С. 524-530.
- [2] Прохоров С.А. Аппроксимативный анализ случайных процессов. – 2-е изд., перераб. и доп. – Самара, СНЦ, РАН, 2001. – 380с.
- [3] Обобщенный спектрально-аналитический метод обработки информационных массивов. Задачи анализа изображений и распознавания образов / Ф.Ф Дедус, С.А. Махортых, М.Н. Устинин, А.Ф. Дедус; Под. общ. ред. Ф.Ф. Дедуса. М.: Машиностроение, 1999. – 357с.
- [4] Батищев В.И. Аппроксимационный подход к оцениванию характеристик взаимосвязи случайных процессов со стационарными приращениями // Изв. вузов. Северо-Кавказский регион. – Технические науки, 2003, №2. С.14-25.
- [5] Батищев В.И., Мелентьев В.С. Цифровые методы измерения интегральных характеристик периодических сигналов: Монография. Самара: Самар.гос.техн.ун-т, 2002. 96с.
- [6] Пиотровский Я. Теория измерений для инженеров. – М.: Мир, 1989. – 335с.
- [7] Батищев В.И., Галкин В.Я., Жуковский Е.Л., Трубин В.А. Методы корректной обработки дифрактограмм // В кн.: Численный анализ методы и алгоритмы. М.: Изд. МГУ, 1986. С.100–111.
- [8] Кременецкий С.Д. Прикладные математические модели: от электродинамики, радиоастрономии... до информатики, телекоммуникации // Зарубежная радиоэлектроника, 2001, №6.

Author index

Индекс авторов

А

- Абрамов Д.В., 246, 260
Алексеев А.В., 270
Андреев В.В., 233, 246, 254, 277
Андреев М.В., 254
Аншаков Г.П., 425
Артиба А., 190
Архангельский С.В., 145
Афонин П.В., 123

Б

- Бадамшин Р.А., 107, 197
Баров В.Г., 18
Батищев В.И., 457
Батищев С.В., 233, 241, 246, 254, 352
Беленький Л.Б., 437
Белкин В.М., 431
Боровик С.Ю., 444, 451
Боровикова О.И., 215
Булгаков С.В., 215, 328

В

- Вайнерман И.А., 203
Виттих В.А., 37, 300, 308, 322, 334,
346, 352, 362
Волхонцев Д.В., 300, 346, 352
Вольман С.И., 270

Г

- Гельфанд М.С., 5
Герасимова И.Б., 107
Горбунов К.Ю., 151
Горлов В.М., 57
Горохов А.В., 370
Гофман И.Д., 176
Гриценко Е.А., 300, 346
Гузаиров М.Б., 107, 197
Гусарова Г.И., 308, 334, 352

Д

- Данильченко В.П., 431
Девятков В.В., 282
Дилигенский Н.В., 84, 96, 334
Донтэн А., 190

Е

- Емельянов В.В., 123, 130
Ефимов А.П., 84

Ж

- Жеребин А.М., 18
Жиганов А.Н., 410, 415

З

- Загорулько Г.Б., 221
Загорулько Ю.А., 215, 328
Захаров В.К., 137
Захарова А.В., 308, 334
Золотухин Ю.Н., 421

И

Иванова Л.А., 322
Ивкушкин К.В., 233, 241, 260
Ильясов Б.Г., 107, 197
Инишев Д.А., 176
Инюцин А.А., 444
Искварина Т.В., 233, 241, 246, 254
Истомин А.Д., 410, 415
Истомина Н.Ю., 410
Истомина С.Н., 166

К

Камаева О.И., 444, 451
Кантемирова Н.И., 357
Карягин Д.В., 246, 260
Кеслер А.Г., 415
Кёних-Риз Б., 203
Кистанов А.М., 209
Коновалов Е.Н., 381
Королева Е.Н., 322
Костов Ю.В., 183
Котельников Г.П., 308, 352
Кропова В.В., 18
Кузин Ю.А., 314
Кузнецов Н.А., 5
Кузнецов С.В., 334
Кузнецов С.И., 352

Л

Лахин О.И., 277, 362
Леонтьев Л.А., 158
Лившиц М.Ю., 386
Липовой Д.А., 183
Любецкая Е.В., 158
Любецкий В.А., 5, 151, 154, 158

М

Мамонтов П.Г., 183
Минаков И.А., 270, 277
Миронов А.А., 163
Мороз А.А., 397
Мостовой Я.А., 425

Н

Назаркина И.М., 375
Напреенко В.Г., 314
Нариньяни А.С., 314
Невзоров В.Н., 170
Невзорова О.А., 170
Нестеров А.А., 421
Никитин А.Н., 346
Никифоров В.А., 241
Носков М.Д., 410, 415

О

Окольнишников В.В., 397
Олейников А.В., 260
Орлов А.Д., 270
Офицеров В.П., 381

П

Павлов В.В., 308, 334, 352

Петров Е.С., 183

Поварова Е.Л., 322

Попов Д.В., 203

Путилов В.А., 370

Р

Райков Б.К., 431, 437

Рапопорт Э.Я., 64

Романов Л.Г., 176

С

Сальков А.Н., 277, 362

Сафронов А.В., 246, 260

Светкина Г.Д., 300, 346

Селиверстов А.В., 154

Секисов Ю.Н., 431, 437

Сидорова Е.В., 215

Симон Н.А., 145

Скобелев О.П., 431, 437

Скобелев П.О., 233, 241, 254, 277,
300, 346, 352, 362

Смирнов С.В., 322, 381

Соллогуб А.В., 425

Солодянников В.В., 386

Сурнин О.Л., 300, 346, 352

Т

Тарасевич В.В., 227, 340, 397

Тарасов В.Б., 113

Тиминский С.В., 294

Томин М.С., 270

Топоров Б.П., 57

Тулупова В.В., 444, 451

Тяпухина Т.В., 380, 334

Ф

Федосеева Л.С., 352, 375

Федунов Б.Е., 403

Федунова Т.И., 403

Х

Хасаев Г.Р., 322

Хасанова Н.В., 107

Холюшкин Ю.П., 215

Ц

Цапенко М.В., 96

Ч

Чернов Е.В., 352

Черняховская Л.Р., 197

Ш

Шалмина Г.Г., 340

Шамашов М.А., 346, 352

Шапиро И.С., 352

Швейкин П.К., 260

Шепитько Ю.О., 130

Шестаков А.А., 370

Ширшин М.А., 158

Шкундина Р.А., 197

Шурбаков А.В., 176

Щ

Щелкунова Л.А., 375

Ю

Юсупова Н.И., 203

Я

Яковлев Н.Е., 314

Ясиновский С.И., 190

Яшин А.Д., 137

С

Chilov N.V., 27

К

Krizhanovsky A.A., 27

L

Levashova T.V., 27

M

Moronski J.S., 45

P

Pashkin M.P., 27

R

Rigby R.H., 76

Rzevski G., 1

S

Shiryayeva O.I., 45

Silbert N.H., 12

Skormin V.A., 45

Smirnov A.V., 27

Summerville D.H., 45

Z

Zarubin V.A., 76

Для заметок

Для заметок

Научное издание

Труды VI Международной конференции

**ПРОБЛЕМЫ УПРАВЛЕНИЯ
И МОДЕЛИРОВАНИЯ В СЛОЖНЫХ СИСТЕМАХ**

14-17 июня 2004, Самара, Россия

Редакторы: Е.А. Федосов, Н.А. Кузнецов, В.А. Виттих

Proceedings of the VI International Conference

**COMPLEX SYSTEMS:
CONTROL AND MODELING PROBLEMS**

June 14-17, 2004, Samara, Russia

Editors: E.A. Fedosov, N.A. Kuznetsov, V.A. Vittikh

Подписано в печать 17.05.2004.

Формат 70x100 /16. Бумага офсетная. Печать офсетная.
Объем 27,9 усл. печ. л. Тираж 500 экз. Заказ 372.

Издательство Самарского научного центра РАН.
г. Самара, Студенческий пер., 3А.

Отпечатано в типографии ООО "ОФОРТ".
443068, Самара, ул. Межевая, 7.
Лицензия ПД 7-0050 от 30.08.2000.
Тел.: 79-08-22, 35-37-01, 35-37-45

