ARL-SR-0418 ● June 2019

DEVCOM
ARMY RESEARCH
LABORATORY

Approaches to Prediction of Cyber Events: Report of the 2017 Specialist Meeting by the North Atlantic Treaty Organization (NATO) Research Group IST-145-RTG

by Dennis McCallam, Tracy Braun, Michael Wunder, Eugene Santos, Teodor Sommestad, Victor Elvira Arregi, Monica Bugallo, Robert Bonneau, Elizabeth Bowman, Mark Mittrick, Marc Jackson, Michael Delucia, Constantin Serban, Angello Sapello, Abhrajit Ghosh, Ritu Chadha, Salvador Llopis, Ignacio Montiel, Juha Kukkola, Juha-Pekka Nikkarila, Mari Ristolainen, Greg Shearer, Nandi Leslie, Paul Ritchey, Frederica Nelson, and Ken Yu

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-SR-0418 ● June 2019

Approaches to Prediction of Cyber Events: Report of the 2017 Specialist Meeting by the North Atlantic Treaty Organization (NATO) Research Group IST-145-RTG

by Dennis McCallam, *Northrop Grumman*

Tracy Braun, Michael Delucia, Greg Shearer, Nandi Leslie, Paul Ritchey, Frederica Nelson, Ken Yu, Elizabeth Bowman, Mark Mittrick, and Marc Jackson, *Computational Information and Science Directorate, CCDC Army Research Laboratory*

Michael Wunder*, Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE)*

Eugene Santos*, Dartmouth College*

Teodor Sommestad*, Swedish Defence Research Agency*

Victor Elvira Arregi*, IMT Lille Douai*

Monica Bugallo, *Stony Brook University*

Robert Bonneau, *USAF Office of Scientific Research*

Constantin Serban, Angello Sapello, Abhrajit Ghosh, and Ritu Chadha, *Vencore Labs*

Salvador Llopis and Ignacio Montiel*, European Defence Agency*

Juha Kukkola*, National Defence University*

Juha-Pekka Nikkarila and Mari Ristolainen, *Finnish Defence Research Agency*

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* June 2019	**2. REPORT TYPE** Special Report	**3. DATES COVERED (From - To)** 21 September 2015–8 October 2017

4. TITLE AND SUBTITLE Approaches to Prediction of Cyber Events: Report of the 2017 Specialist Meeting by the North Atlantic Treaty Organization (NATO) Research Group IST-145-RTG	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Dennis McCallam, Tracy Braun, Michael Wunder, Eugene Santos, Teodor Sommestad, Victor Elvira Arregi, Monica Bugallo, Elizabeth Bowman, Mark Mittrick, Marc Jackson , Michael Delucia, Constantin Serban, Angello Sapello, Abhrajit Ghosh, Ritu Chadha, Salvador Llopis, Juha Kukkola, Juha-Pekka Nikkarila, Mari Ristolainen, Greg Shearer, Nandi Leslie, Paul Ritchey, Frederica Nelson, and Ken Yu	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US CCDC Army Research Laboratory (ATTN: FCDD-RLC-HC) Aberdeen Proving Ground, MD 21005	**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-SR-0418
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) NATO Science and Technology Organisation Collaboration Support Office (CSO) BP 25, 92201 Neuilly sur Seine, France	**10. SPONSOR/MONITOR'S ACRONYM(S)** NATO
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Traditional approaches for gaining cyber domain situational awareness seldom consider factors associated with the adversary's capabilities and behaviors. The 2017 NATO Specialist Meeting, IST-145, on Predictive Analytics and Analysis in the Cyber Domain intends to show that if these factors are taken into consideration, multiple and cooperative analytic approaches can predict exploitation of known vulnerabilities even if the attack pattern is previously unknown. Furthermore, we intend to show that such predictions provide meaningful temporal mission impact alerts to operators and commanders, and can move cyber defense from reactive to proactive. This will help maintain NATO and national security.

15. SUBJECT TERMS

predictive analysis, predictive analytics, adversarial cyber operations, cyber situational awareness, and cyber defense

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Tracy Braun
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER (Include area code)**
Unclassified	Unclassified	Unclassified	UU	247	(301) 394-4954

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Foreword

This foreword provides an explanation and differentiation of NATO IST-145 and IST-129. In this report, both IST-145 and IST-129 are referenced.

In 2013, the NATO, Science and Technology Organization (STO), Information Systems Technology (IST) panel approved a Technical Activity Proposal (TAP) for IST-129, Predictive Analysis of Adversarial Cyber Operations. The approval of a TAP means a Research Task Group (RTG) is formed and member countries begin collaborating on the research topic. One of the tasks for IST-129 was to sponsor a collaborative event for predictive analytics.

Simultaneously, the NATO STO IST panel approved a TAP for a Research Specialist Meeting (RSM) for Predictive Analytics, IST-145, in 2015. The intent was to schedule a Specialist Meeting on the generic topic of predictive analytics in the 2017 timeframe. To expedite and consolidate things, the IST-129 RTG assumed the lead for IST-145 RSM to achieve both goals. The Specialist Meeting in October 2017, the topic of this report, was the capstone event for IST-145. When referencing items specific to the Specialist Meeting, or IST-145 research, the authors tried to use IST-145. When referencing research performed as part of IST-129, the authors used IST-129. In some cases, both are referenced for accuracy as the two activities had similar and overlapping goals and expectations.

Tracy Braun, CCDC Army Research Laboratory

1. Introduction

This report summarizes the discussions and findings of the 2017 NATO Specialist Meeting, IST-145, on Predictive Analytics and Analysis in the Cyber Domain. The Specialist Meeting was held in Sibiu, Romania, on 10–11 October 2017 at the Nicolae Bălcescu Land Forces Military Academy. The Specialist Meeting chairman was Dr Dennis McCallam, Northrop Grumman, United States. This workshop was unclassified and open to NATO nations, Partner for Peace nations, Mediterranean Dialogue, Istanbul Cooperation Initiative nations, and global partners.

The Specialist Meeting committee was composed of David Aspinall, University of Edinburgh, United Kingdom; Tracy Braun, US Army Research Laboratory (ARL)*, United States; Roman Faganel, Slovenia Ministry of Defence, Slovenia; Leonard Ferrari, US Naval Postgraduate School, United States; Heiko Guenther, Fraunhofer FKIE, Germany; Matthew Kellet, Defence R&D Canada, Canada; Joseph LoPiccolo, US Naval Postgraduate School, United States; Peeter Lorents, Estonian Business School, Estonia; Wim Mees, Royal Military Academy, Belgium; Juha-Pekka Nikkarila, Finnish Defence Research Agency, Finland; Teodor Sommestad, Swedish Defence Research Agency FOI, Sweden; and Margaret Varga, Seetru Ltd. and Oxford University, United Kingdom.

In the organization and planning for this Specialist Meeting, we examined and analyzed the current state of practice with respect to prediction of cyber behavior and areas that could contribute to the ability to effectively predict adversarial cyber behavior at some level.

At the outset and in the planning phases, we developed four primary objectives:

1) Bring together in one group or forum, subject-matter experts researching and developing predictive analytics/analysis (PA) tools for use with big data (hard and soft) in order to improve understanding and share thoughts on predictive analytics;

2) Bring together researchers, practitioners, and vendors to discuss the state of the art and practice on PA in the cyber domain;

3) Provide a forum to present current tangible and theoretical research in the field of PA of adversarial cyber operations; and

* The work outlined in this report was performed while the US Army Research Laboratory (ARL) was part of the US Army Research, Development, and Engineering Command (RDECOM). As of 31 January 2019, the organization is now part of the US Army Combat Capabilities Development Command (formerly RDECOM) and is now called CCDC Army Research Laboratory.

4) Investigate and suggest an international way forward to progress the state of the art and implementation of adversarial cyber behavior prediction.

To accomplish these objectives, we sought papers, research, use-case studies, and/or analyses on the PA of adversarial cyber operations covering a wide range of topics:

- Predictive tools being used in big data
- Findings or experiments on relationships between algorithm types implementing analytics and domain of implementation
- Fusion of different analytic approaches for prediction of nonsignature-based cyberattacks
- Cyber situational awareness conveyance tools and methods situation description
- Detection of threat capability, and course of action (COA) selection as a function of threat capability, as defined by the Defence Science Board (DSB)
- Evaluation of threats that leverage known vulnerabilities with previously unseen exploits
- Characterization of adversarial behavior within a network including tactics, techniques, and procedures (TTPs)
- Methods for detecting unknown vulnerabilities
- Measures and metrics of adversarial cyber activity
- Methods for dealing with adversarial adaptation to predictive models
- The cyber observe–orient–decide–act (OODA) loop

2. Background

While the growth of available data has increased exponentially, the capabilities of analysis tools, recognition software, and computer capacity have not grown nearly as fast, though they are still far more powerful today than even a decade ago. Several PA tools that are in the early stages of research show great promise for improving our understanding and ability to support decision making at reduced levels of risk. At the same time, the challenges of the 21st century have also become more complex and include the impact of a volatile global economy, population migrations, changing weather patterns due to climate change, loss of arable land

and fresh water on a global scale, expected population growth, pandemics, and terrorist activities worldwide. Having a good indication of likely future actions by nation states, terrorist organizations, refugees, and financial markets has become vital to the planning of collaborative organizations such as NATO in order to form improved preventative and response strategies to potential large-scale crisis events. The PA tools available to analysts today are quite powerful when compared to those of a decade ago. The problems that can be supported by PA range from Commanding Officer decision support in peacekeeping and conflict zones, to strategic decisions based on future global requirements and regional support needs due to predicted pandemic and other health issues, to prediction of natural disasters needing high-availability disaster recovery (HADR), to detection of anomalies on critical communication and control data networks, that is, cybersecurity. Some of the required predictions need to be used in decision making in real time or even within microseconds of an occurring event, while others can be more strategic and even utilize massive offline computation. The variables associated with these major challenge areas has led to the development of a collection of PA tools and research programs with differing properties. There are already a number of tools that are being developed to provide predictions from the rapidly growing available world databases, but often there is little crosstalk among researchers developing some of the most effective predictive tools.

There exist approaches (e.g., Brown et al. [2002] and Kott and McEneaney [2006]) to the PA of adversarial COAs in noncyber domains, although the efficacy and robustness of these approaches remains uncertain. The shift of military operations to a reliance on cyberspace over the last 25 years and the speed of actions in that domain lead to a need to be proactive in understanding how attacks happen and, more importantly, what is likely to occur in the future as a result.

PA has been widely relied upon to evaluate options in many domains such as banking, gaming, insurance, and retail. These techniques have not yet been applied to the cyber domain, likely because there are significant challenges in doing so:

- Cyberspace is complex, dynamic, asymmetric, and not well understood, making the adversary’s choice of potential attack steps much larger than in other domains.

- The adversary has the upper hand because their actions in cyberspace are much less observable and take less time than in other domains.

- The rapid evolution of new zero-day exploits obscures (full situational awareness) knowledge and temporal awareness of the current situation.

- There are diverse cultural, social, and cognitive traits of the adversary that are important factors in determining future adversarial COAs.
- Coordination among nations and transnational institutions requires close collaboration to enable extremely fast exchange of knowledge about adversaries and their anticipated operations using a common set of concepts, terms, and methodologies.

There are aspects of adversarial actions and the cyber domain that can be used to our advantage in PA. It may be possible turn the temporal advantage of the adversary's quickness of action to our advantage if we can get inside of their decision-making (or OODA loop) cycle to make timely and accurate predictions of their future actions. We can also use our knowledge of the adversary's capabilities, and the maturity thereof, to reduce the space of possible adversarial actions and increase the accuracy of our predictions (Linkov et al. 2013). In fact, a DSB report (*Resilient Military Systems and the Advanced Cyber Threat*) focuses on assessing capabilities and analyzing the specific tools and TTPs used by the threats.

The purpose of this Specialist Meeting was twofold. The first was to look at the science of PA in general and the second to consider implementations of PA specifically with regard to predicting adversarial cyber operations.

3. Detailed Review of the Presentations

This Specialist Meeting explored how the directions of current and future science and technology may impact and define potential breakthroughs in the field of prediction as applied to the cyber domain. The presentations and discussions, along with relevant committee conclusions at the Specialist Meeting, are contained in this report. This section of the report summarizes each presentation and then provides a set of committee observations and conclusions.

3.1 Introduction to the Specialist Meeting

Presented by: Michael Wunder, NATO Information Systems Technology (IST) Panel Chairman

Dr Wunder provided a welcome and official kickoff for the Specialist Meeting. His presentation provided the background of how the Specialist Meeting activities fit into the NATO research scheme. He did a very high-level review of the Science and Technology Organization (STO), which now consists of the Collaboration Support Office, the Office of the Chief Scientist, and the Centre for Maritime Research. STO has a well-defined charter as the strategic enabler developing technology advantages for defense. STO helps to promote different science and technology (S&T) activities, not only enabling but influencing the defense capabilities and threat mitigation, and supporting NATO decision makers. The salient feature of STO is twofold. First, STO helps organize activities that are common S&T problems across the alliance, which in turn create additional relationships between and among researchers from member nations. This accelerates the trust necessary for research cooperation. The second feature may be the most important, since STO activities help force multiply investments individual countries make on projects through common research on common problems. The impact to the work of IST-145 and IST-129 is that the issue of cyberspace adversary prediction is a universal problem and coming together to examine collaboration potential is of keen interest.

Dr Wunder went on to describe the panels and groups chartered by and supporting STO. There are seven panels: Applied Vehicle Technology (AVT), Human Factors & Medicine (HFM), Information Systems Technology (IST, and the oversight Panel for IST-145 and IST-129), Systems Analysis & Studies (SAS), Systems Concepts & Integration (SCI), and Sensors & Electronics Technologies (SET); and one group, Modelling and Simulation Group (MSG). He noted that STO is encouraging cross-panel cooperation and there has been an uptick in cosponsored activities. The IST Panel oversees the cooperation that results in systems improvements with a focus on cybersecurity and secure information transfers. It comprises 54 members representing 45 countries and associates. IST sponsors three focus groups: Decision Support, Ensuring Communications, and Security & Trust. He also reviewed the six ways of participation:

- Exploratory Teams (ETs) assist or advise the panel on the technical merit or feasibility of a specific longer-range proposal for a technical activity or future content of the Panel's technical program.

- A Research Task Group (RTG) is chartered for a maximum of three years to address and provide documentation against a particular and specific research and technology problem.

- A Research Symposium (RSY) promotes the exchange of state-of-the-art knowledge among a wide audience on an important scientific or applied topic.

- Symposia, Specialist Meetings, and Workshops aimed at promoting exchange of state-of-the-art knowledge and facilitating intensive information exchange and focused discussion among an audience of invited specialists and keynote speakers.

- Lecture Series and Technical Courses aimed at disseminating state-of-the-art scientific knowledge and recent field developments through onsite instructor training to meet the needs of NATO.

Dr Wunder concluded by citing some specific examples of current work and to inform the group on an upcoming Specialist Meeting on Big Data and Artificial Intelligence, IST-600, to be held in Bordeaux, France, 31 May–1 June 2018.

3.2 Setting the Stage: A Review of the Work of IST-129, Predictive Analysis of Adversarial Cyber Operations

Presented by: Dennis McCallam, Specialist Meeting Chairman

Dr McCallam, as the chair for the sponsoring activity IST-129, gave a review of IST-129 work to date to provide the technical context for the Specialist Meeting. In essence, PA has to consider the past and present to predict the future. The current IST-129 group has been sponsored by nine member nations and is into the second year of the three-year remit of work. The research task group has three objectives:

- To characterize the current state of research in the field of PA of adversarial cyber operations. This will be satisfied through an assessment of approaches concentrating on cyber battlefield intelligence preparation, describe the similarities and differences with conventional warfare approaches with respect to PA of adversarial COAs, and validate the current state of the art through a workshop activity.

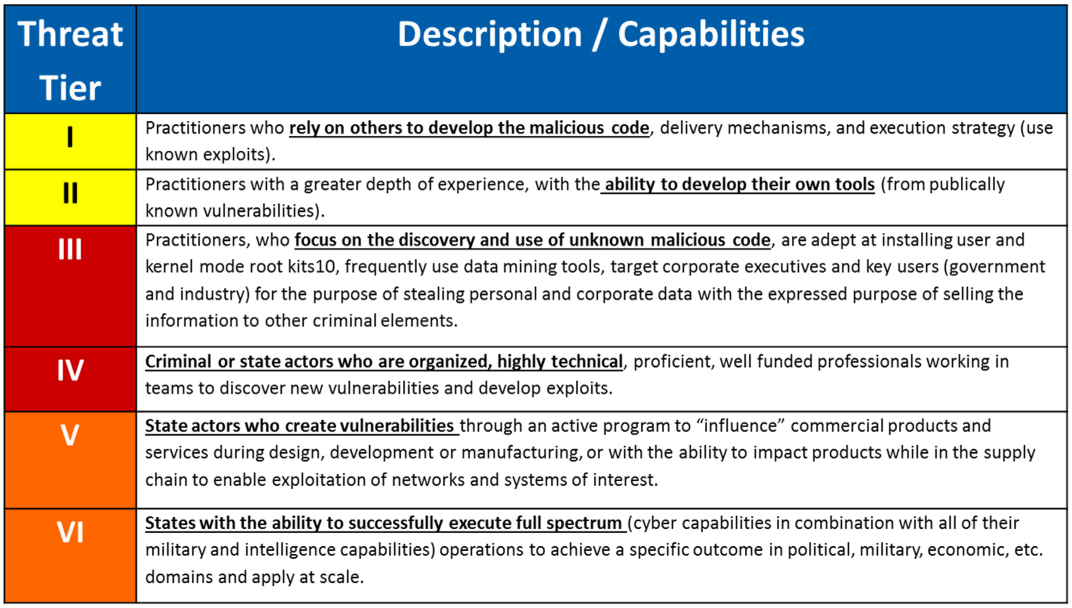
- To develop an initial roadmap for development of a comprehensive set of methodologies, technologies, and tools for advancing the proactive PA of adversarial cyber operations.

- To develop a final technical report that supports NATO and its members.

To date, we have discovered very little work in this area at least in the unclassified domain. As an example, the IST-129 committee (in preparation for this Specialist Meeting) contacted over 100 companies, and most felt their technology readiness level (TRL) in any solution was not high enough at that time.

Early on in the research, the committee established some key ground rules and the most important of those concerned information. The committee decided that all information used in the work of the committee would be unclassified and open source. Also the committee felt that some noncyber areas look at machine learning (ML) and data mining so there was potential in evaluating some of these areas.

Initially, the committee selected a definition of the threat (Fig. 1) as found in the 2013 US DSB, *Resilient Military Systems and the Advanced Cyber Threat* (Linkov et al. 2013), which defined cyber threats in terms of capabilities as opposed to identifying specific groups. This allowed the work of the committee to address threats in terms of capabilities, which is universal in terms of the cyber threat but avoids potential classification issues of specific group identification. This capability description has six levels organized into three bands of capabilities. Levels I and II concentrate on threats that leverage known vulnerabilities using known exploits. Levels III and IV concentrate on threats that focus on known vulnerabilities using unknown exploits. Levels V and VI are more the state actors that have the capabilities to create unknown vulnerabilities and associated unknown exploits. From a financial investment point of view, operating at levels I and II is very cheap. The investment in capability development escalates with levels V and VI, which are very expensive. From a focus area, the committee eliminated levels I and II, since these are deterministic areas that are addressed through signature detection. The committee elected to not "boil the ocean", so elected to focus the activities on level III.



Threat Tier	Description / Capabilities
I	Practitioners who **rely on others to develop the malicious code**, delivery mechanisms, and execution strategy (use known exploits).
II	Practitioners with a greater depth of experience, with the **ability to develop their own tools** (from publically known vulnerabilities).
III	Practitioners, who **focus on the discovery and use of unknown malicious code**, are adept at installing user and kernel mode root kits10, frequently use data mining tools, target corporate executives and key users (government and industry) for the purpose of stealing personal and corporate data with the expressed purpose of selling the information to other criminal elements.
IV	**Criminal or state actors who are organized, highly technical**, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.
V	**State actors who create vulnerabilities** through an active program to "influence" commercial products and services during design, development or manufacturing, or with the ability to impact products while in the supply chain to enable exploitation of networks and systems of interest.
VI	**States with the ability to successfully execute full spectrum** (cyber capabilities in combination with all of their military and intelligence capabilities) operations to achieve a specific outcome in political, military, economic, etc. domains and apply at scale.

Fig. 1 Description of cyber threats with respect to their capabilities. Derived from page 22 of the DSB report, *Resilient Military Systems and the Advanced Cyber Threat* (Linkov et al. 2013).

The committee also examined the cyber kill chain with respect to predictive countermoves that would essentially move the threat into a constant reconnaissance position as opposed to positions in the kill chain that could be viewed as more dangerous. The committee examined Boyd's OODA loop, originally designed for fighter pilots, to analyze how the OODA loop could be purposed in the cyber domain. The "cyberization" of the OODA focused on the orient phase utilizing cyber event DNA, identification as to where someone may have learned the craft, new cyber-related information that augments the previous experiences and analyzes phases. There was also considered the notion within the OODA loop that there could be multiple outcomes including an optimized prediction, an interim prediction or the identification of data shortages during the decision phase. The latter implies a valid action could be to seek additional and/or specific data.

At this point in the research agenda, there have been several interim conclusions. First, the known vulnerability/known exploit is a solvable problem and has been solved, but not necessarily implemented. Since this is a signature identification problem, it is more of a detection identification problem as opposed to prediction. Prediction using pattern matching is trivial. Second, the prediction edge values (0% and 100% certainty) are unattainable. The committee felt this because the next cyber incidents are not necessarily dependent on previous cyber events, but rather a more independent variable. The example given here was taken from lottery games where number with highest frequency of occurrence are often displayed, which gives the illusion that the next lottery draw is a function of the previous draw.

Third, it appears inclusion of feedback earlier and in multiple areas of OODA will enhance/streamline prediction and the committee considered this a topic for future research. Finally, identifying the attacker capabilities as a function of the DSB tiers is hard at the beginning of the analysis, which is characterization of the attacker. Methodologies for levels I and II are more certain than levels III, IV, V, and VI with the implication that the methodologies for levels V and VI are different than for levels III and IV. The implication here is effective threat analysis using the DSB criteria appears to infer there are three different processing/analysis approaches each based on threat capability.

At the time of this Specialist Meeting, there are some interim conclusions from the work of the IST-129 committee thus far:

- The known vulnerability/known exploit is a solvable problem and has been solved, but not necessarily implemented through automation. It is detection as opposed to prediction, making prediction in this case trivial.
- Prediction at the edge cases is outside scope of effective prediction at this time. The edge values on the "known vulnerability – unknown exploit" capability threat (0% chance a cyber event will not occur and 100% certainty that a cyber event will occur) are potentially unattainable:
    - Incidents can be independent variables and can have no relation to previous cyber events. There is no guarantee that the sequence of

    cyber events identified represent a fully understood and known threat TTPs.

    - A prime example of this in real life are lottery games that present the occurrences of numbers in the previous *n* draws tricking people into thinking the next draw is a function of previous draw(s).

    - The IST Task Group felt that the Colin Powell credited quote—"As an intelligence officer, your responsibility is to tell me what you know. Tell me what you don't know. Then you're allowed to tell me what you think. But you always keep those three separated."—has importance in the prediction process.

- The IST-129 Task Group felt that a common taxonomy was needed to communicate in the cyber prediction domain and recommends the use of Structured Threat Information Expression (STIX) as a consistent means of enhancing communication.

- Inclusion of feedback earlier in a "cyber OODA loop" appears to enhance/streamline prediction, which is a potential topic for future research. This could infer that a next step in prediction could involve correction in a manner similar to Kalman filtering. One constraining issue identified is the temporal dimension and the need to process in real-time efficiency.

- Discerning which capability tier within the DSB framework to characterize an attacker is hard at the beginning of the analysis. For example, methodologies for identifying attackers with capabilities defined in levels I and II (known vulnerabilities – known exploits) are completely deterministic and more precise and defined than attackers in capability levels III, IV, V, and VI.

- Related to the previous comment, the task group notionally agreed that there are unique methodologies for identifying and predicting threats at different levels within the DSB threat capability definition. The implication for practitioners is that for each threat capability family (known vulnerabilities – known exploits ; known vulnerabilities – unknown exploits; and unknown vulnerabilities – unknown exploits) each processing stream is different, further supporting the notion that one algorithm does not solve the threat identification or prediction problem.

3.3 Keynote Presentation: Adversary Intent Inferencing for Predictive Analytics

Author: Dr Eugene Santos, Thayer School of Engineering, Dartmouth College

In performing some of the committee analysis into the area of PA, there was one researcher who had done substantial work in the PA area. The Specialist Meeting was fortunate to have Dr Santos as the keynote speaker. His topic was adversary intent inferencing for PA. The focus of the keynote was on determining adversary intentions and understanding what drives those actions. The domains of discussion are on military operation, planning, and intelligence analysis.

One reason modeling adversaries is difficult is the level of uncertainty in predictions and the relatively wide-open nature of research in this space. Intent inference, or user intent inference, involves deducing an entity's goals based on observations of that entity's actions (Geddes 1986). In turn, this becomes useful for generation of advice and the definition of future information requirements (Bell et al. 2005; Santos 2003). There are some approaches to intent inferencing:

- Plan-goal-graph (PGG): a network of plans and goals, where each high level goal is decomposed into a set of plans for achieving it, and the plans are decomposed into subgoals, which in turn are decomposed into lower-level plans (Geddes 1994). Intent is finding the path from observables to a plan or goal.
- Operator function model (OFM): an expert system using a heterarchic-hierarchic network of finite-state automata, in which nodes represent entity's activities and arcs represent conditions that initiate/terminate certain activities (Rubin et al. 1988a, 1988b; Bushman et al. 1993; Chu et al. 1995). Connect observed action to appropriate activity trees.
- Generalized plan recognition (GPR): this recognizes the entity's plan for carrying out the task, based on observations, an exhaustive set of discrete actions (a plan library), and constraints (Carberry 1988; Goodman and Litman 1990; Lesh et al. 1998).

Intent becomes important because it can help one predict the future, explain the present, and understand the past. Additionally, understanding and identification of intent can help prune the search space, bound optimization, guide scheduling, and better allocate resources.

Traditionally, Blue (friendly forces) COAs were wargamed against the "most likely/dangerous" Red (adversary) COAs (circa 2001), but these were more often pre-scripted as opposed to being more dynamic. Asymmetry of capabilities and asymmetric threats both mean differences in intent. The question becomes more of an issue of how one does assessments or what-if analyses.

Essentially the goal is to develop better adversarial modeling. This spawns the question of identifying what one needs to know about the adversary. Intent is not just a plan or an enemy COA, but also considers the why. Some of this can be ascertained by looking at what will happen next. The definition is Adversarial Intent = Goals + Beliefs + Actions + Commitment. Adversarial modeling becomes useful in financial/business competition (game theory), politics/elections, sports, and so on.

This should be straightforward since evaluating the goals the enemy can be defined as enemy goals = pursuing + the support of those goals + the plan to achieve it. To understand and predict Red COA, one needs to model from the enemies' perception (point of view [POV]). This avoids accidentally imposing Blue beliefs on Red and also allows modeling of deception.

The US Air Force Adversary Intent Inferencing (2001–2004) program examined Effects-Based Operations (EBO), influencing enemy COAs by carefully selecting and executing our own COAs to achieve desired objectives. How we determine those EBOs is based on three formative components (enemy foci, enemy COA, and enemy rationale). The core adversary intent model contains three components: 1) goals/foci of what the adversary is doing, 2) rationale network or why the adversary is behaving that way, and 3) action network or how the adversary is achieving its goals.

Looking at this from a static behavior point of view, the enemy observables were fed into Bayesian networks for enemy rational and enemy actions. This included adversarial axioms (X), adversarial beliefs (B), adversarial goals (G), and adversarial actions (A) to avoid the infinite regression scenario (e.g., "I know that he knows that I know that he knows…").

Next, dynamic behavior models were examined. Dynamic behavior, or emergent adversarial behavior, shows how an adversary changes over time. Missions differ based on different intent. An example scenario from Nellis Air Force base was reviewed. The scenario consisted of two different commanders with two different tactics. One commander was aggressive and the other was passive. The concept of Bayesian knowledge fragments was introduced, which estimated enemy intentions based on sequence of Red–Blue interactions such as depletion of resources. In the simulation, the aggressive commander was more likely to actively respond. The passive commander had higher likelihood to defend and conserve resources (such as ammunition). Counterintuitively, the passive commander caused more damage in the scenarios and preserved more assets by shutting down, and making their forces harder to target. Over time, the aggressive commander used up their ammunition and could no longer inflict damage. The passive commander could still fire at enemy forces while retreating or returning home.

Dr Santos introduced the concept of the dynamic adversarial gaming algorithm (DAGA). DAGA develops algorithmic techniques to accurately predict community of interest (COI) responses to social, cultural, political, and economic actions. It incorporated various learning aspects: each different play has different outcome. It gives one a graph of possibilities. Cultural differences were shown to be important with respect to the gaming. What does one need to know about the adversary? What is rational? These questions were based on social, cultural, economic, and political parameters.

It also allows for Bayesian fusion of these factors to model different groups, in different conditions, to make them more asymmetric in simulations. To highlight DAGA's capabilities, it was integrated with the popular Civilization 4 (2005–2008) game engine to demonstrate how the infusion of sociocultural influences lead to a much more realistic asymmetric adversary.

Next Dr Santos talked about his most recent work modeling complex adversaries and their intent. This work uses a networked intent model, with evolving behaviors, for multiple adversaries. The goal of this work is to help commanders and decision makers by modeling targets as complex, adaptive systems. The model can produce timely, correct, and actionable intelligence for the warfighter, when the system has only partial observable assets, fluid environments, multientity situations with dynamic friends, foes, and neutral parties. He used an example of a Somali pirate group, where the structure of the group was modeled as a network hierarchy, with different roles, lines of communication, and social ties.

His future work includes plans for learning adversary intent using dynamic decision models.

Conclusions: Dr Santos has been working in the field of adversarial modeling and predicting adversarial intent for many years, and is an expert in this field. He has some of the most dynamic, complex models of adversarial intent the panel could find in literature searches. The inclusion of social and cultural factors will be important as the models get more complex and more accurate. The development Dr Santos has conducted in the methodology of modeling adversaries could be of great benefit when modeling the adversarial intentions in the cyber domain.

3.4 Position Paper

Author: Teodor Sommestad, IST-129 RTG

Dr Sommestad provided an overview of the IST-129 RTG's findings on PA of adversarial cyber operations. The IST-129 RTG's position paper was the catalyst for the formation of the IST-145 Specialist Meeting on Predictive Analytics and Analysis in the Cyber Domain.

The position paper represented the RTG's survey of the state-of-the-art and current research in PA. The RTG developed a roadmap and used it to guide the development of our final technical report. We also more clearly defined the problem and identified issues with using PA with cyberspace.

When the RTG first met, Dr Alexander Kott had proposed the problem of prediction as a closed-loop control problem. Inputs to a predictive controller must be fed into a model, and predicted outputs must be fed back as inputs to make adjustments for future predictions. However, the RTG had a hard time finding examples of when this is done in cybersecurity in this way or in other security domains. There are many examples of when one makes predictions and influence the "process", but one rarely has an explicit model of how a certain future input would alter the process.

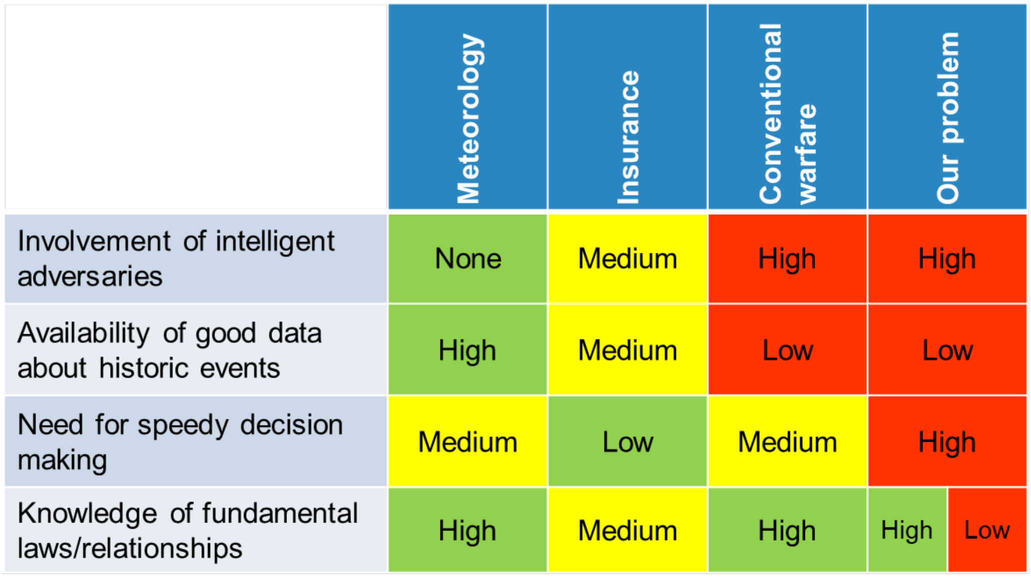
To make useful predictions, one also needs accurate and up-to-date situational awareness. This is a difficult problem in dynamic systems in cyberspace. People make projections today. There is little tool support for making predictions, and much is guesswork. There is no table where one can look up every situation and simply determine what one should do. However, there are in other domains (e.g., in air traffic control) that have plenty of manuals and regulations.

In information systems, people say that there are five types of theories: analysis, explanation, prediction, explanation and prediction, and design and action. The RTG is concerned with making predictions. Clearly, to make predictions about upcoming cyberattacks would be nice. Thus, we would want theories to make predictions. But, if we also want predictions to help a system administrator, we probably want some causal explanation that the administrator can assess the reasonableness of. So, looking at it this way, the RTG was set up with the following:

- Looking for a very powerful theory over the cybersecurity domain,
- Identifying what intelligence it would need, and
- Saying how the cybersecurity domain should go about to develop methods and tools to create this theory.

That should not be impossible; people make predictions in other domains.

However, the RTG is in a tough position. Compared to other domains where predictions are made, we have intelligent adversaries, poor data, a need to make quick decisions, and a poor understanding of the fundamental laws or relationships of our system. Figure 2 presents the characteristics of predictions in various domains.



	Meteorology	Insurance	Conventional warfare	Our problem
Involvement of intelligent adversaries	None	Medium	High	High
Availability of good data about historic events	High	Medium	Low	Low
Need for speedy decision making	Medium	Low	Medium	High
Knowledge of fundamental laws/relationships	High	Medium	High	High Low

Fig. 2 Characteristics of predictions in various problem domains

When it comes to the fundamental science and laws of our problem, we have an advantage in the sense that the fundamental laws are manmade (e.g., binary code) and can be documented. However, we have a disadvantage in the sense that many attacks actually compromise the laws/relationships we think exist (e.g., by using zero-day attacks or the memory of computers in ways we did not want or anticipate).

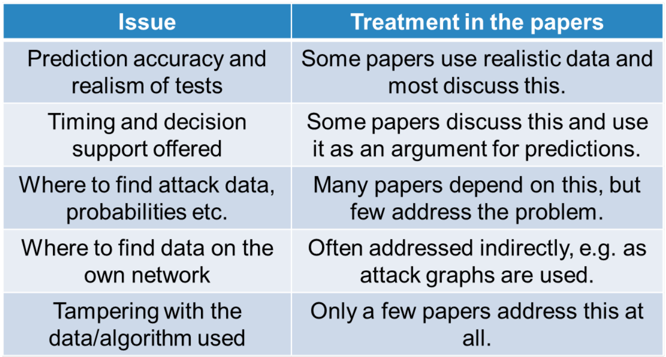
For our literature search, the RTG focused on research describing a solution explicitly developed for predicting adversarial cyber operations. Predictions should be a statement about what will the future, for example, “attack XYZ will be the next one directed toward us”, “the probability we are attacked with XYZ is 17%”, or “the probability of the attack XYZ is 17% in the coming year”. Assessing what is possible does not suffice for our purposes. The solution should be described in a reviewed paper.

For our literature search, it was a collaborative effort. We performed several types of searchers, systematic searches, ad-hoc searches, and searched the citations in relevant papers. We found 35 related papers. About 7–15 of these papers met our criteria.

As the original IST-129 statement of work included threat intelligence; therefore, we wanted to look at the information requirements of different methods. We started classifying their input in terms of STIX, a Mitre standard for exchanging threat information. It was more difficult than expected, and there were so few papers in our final database, so we did not finish this classification. We shifted to a more qualitative review of the relevant papers.

We looked for how the papers dealt with relevant issues (Table 1). Far from all papers address these issues. For example, few papers start with an analysis of how fast one needs to respond or what confidence levels one needs to dare to make decisions. A problem with antagonists are that they can attack the prediction mechanism, and fool us. This is hardly discussed at all.

Table 1 Relevant issues related to literature survey



Issue	Treatment in the papers
Prediction accuracy and realism of tests	Some papers use realistic data and most discuss this.
Timing and decision support offered	Some papers discuss this and use it as an argument for predictions.
Where to find attack data, probabilities etc.	Many papers depend on this, but few address the problem.
Where to find data on the own network	Often addressed indirectly, e.g. as attack graphs are used.
Tampering with the data/algorithm used	Only a few papers address this at all.

Some conclusions from our literature search were as follows:

- Predictions based on analogy or pattern matching are common (e.g., in antivirus systems).
- Predictions based on a generic model are few.
- Threat data are scarce, have quality issues, and can be “attacked”.
- Attacks tend to break the rules and laws we set up, or think we have.
- Plan recognition is used, not models over adversary intentions.

3.5 Efficient Monte Carlo Methods for Prediction in High-Dimensional Systems with Big Data

Authors: Victor Elvira Arregi and Monica Bugallo

This approach was described using filtering/prediction and then model learning. Interesting problem that was approached as a four-step process: first perform filtering to estimate the current state, next predict the future state, then predict the future observation, and finally smooth the past state estimate. This produced a distribution (via Monte Carlo) of outcomes all with attached probabilities of uncertainties.

The focus is on dynamical models and compare to bioinformatics, geographical information systems (GIS), and imaging, which are closer to the cyber problem of predicting behavior in the known and unknown domains.

The methodology used seems mathematically rigorous and since the functions are known it is likely to give good estimations of the current state. The model discussed predicts the future observation by applying statistical Bayesian approach. They address this by random measuring (Monte Carlo) and calculation for analysis for any model—this is important, since this is coupled with the Bayesian approach it notionally implies multiple algorithmic approaches that include the measure of uncertainty. The next step was to measure the uncertainty of prediction by applying statistical Bayesian approach—another important observation.

The conclusions were the following:

- Different state space models (SSMs) require different number of samples for operating at the same level of accuracy (even the same SSM at different states). In addition, recent advances in multiple importance sampling (MIS) and adaptive importance sampling (AIS) allow to use few samples and still have a great performance.
- When one applies sampling, one always has to consider the risk of whether the sampling is good or not (if the distribution of the sampling is actually something different that one assumed). Some level of intuition is necessary: if the filter is not applicable the predictions are biased. In some ways, this is related to both the observation and orientation phases of Boyd's OODA loop.
- The intent is to mathematically prove how to filter data based on the historical data of observations, which appears to correlate to Known, Known and Known, Unknown problems, in our opinion.

- The particle filtering approach, in looking “backward” in time, appears to be an inverse Kalman filter. There is also the notion that particle filtering will fail in high-dimensional systems.
- The four-step process, in the opinion of the committee, appears to directly address potential hypotheses generation that could be used in a cyber domain.
- The conclusions with respect to big data in the SSM reviewed support the notion that we could predict the occurrence of a set of given events but not what the next event will be.
- In dynamically predictive context, we must approximate the evaluation for each sample and reprocess the entire set of observables indicating this appears to be a more computationally intensive approach.
- This is a two-algorithm approach with one checking uniformity of the model with statistically proven method (Bayesian) and the other checking the autocorrelation (Monte Carlo). This appears to support the notion, in the committee’s view, that correlation in cyberspace will not use single-algorithmic approaches.

3.6 Predicting Adversarial Group Membership and Activity in Cyberspace

Authors: Elizabeth Bowman, Mark Mittrick, and Marc Jackson

This discussion was centered on social understanding and reasoning framework (SURF) tool development, which was a ARL-funded TRL-6 activity installed at Aberdeen Proving Ground, Maryland. Too much data is a common complaint in most operational domains and this limitation requires decision makers to mentally reconstruct, infer, and extract relevant information through laborious and error-prone internal processes. The paper addresses the need for the timely extraction and prioritization of high-value, decision-relevant information. The expanding cache of interesting data is a common complaint in most operational domains. There is an increasing complexity of military challenges; roughly 2.5 billion persona are social media users, complicating the problem for intelligent operators.

SURF finds and fingerprints social media users based on interactions and was applied to “ISIS”, “Business”, and Hacker classes (Twitter). Fundamental questions addressed were, Who is important in this adversary network and how are they influencing people? As the scenario example, they used Twitter messages to find adversaries (who is important in the adversary network) and one application used the density of communication of ISIS sympathizers to categorize the Twitter users in order to find the leaders.

The following are the conclusions:

- This produces two outcomes: 1) a list of potential ISIS affiliates and those most important within the network of potential ISIS affiliates, and 2) a list of their influencers.
- Initial testing results indicate that there is a savings of 80%–85% in analytic processing time over current analyst approaches.
- Analysts can create tailored watch lists based on the social networks of those classified as likely ISIS members.
- With this method one may find the interesting social accounts. Finding the actual adversaries in person is an open question that was not asked.
- Practical approach for analyzing relationships across a known group through analyzing social media interactions. They evaluate edge cases (followers in this scenario) to "predict" if they are members of the known set (in this case ISIS members).
- Given that this approach analyzes relationship entities in text to identify potential members of a threat group, the committee agreed that this approach for this use-case is useful in clarifying relationships. Potential application in cyber domain is not so much prediction, but rather given a set of cyber events (the "messages" from this paper) what could be hierarchy or the relationship across those events.
- Committee noticed that there is really not a temporal (time) dimension within this specific use case and wondered what impact that could have when this approach was ported to a cyber domain.
- Another open question was the possibility of this approach being applied to insider detection.
- Filter data with ordered ranking based on eigenvector centrality of each node result in a size-ordered circular layout. One should label graphs with ego notes to identify the most influential personas.

3.7 Shaping Cyberspace: Data and Decision Analytics

Author: Robert Bonneau

The author proposes to form a framework to analyze big data. There is currently no standard way to implement and assess performance for data analytics. Current approaches have heterogeneous data sources/algorithms without ground truth making it hard to know what capability is being purchased along with few performance measures. This presentation is more about standardizing the data structure and the representations then prediction.

He suggests a new approach where the analyst is not in the middle of information loop but rather sits on the loop increasing quality of service. The vision is for a cloud-based approach that is based on an open-standard approach that reconfigures known and trusted components to satisfy multiple missions.

Standard threat or mission graphs and the associated data needed to assess a particular threat are can be available for baseline assessment and design of future mission analysis.

As an example, UAVs had a lot of radio interference and were losing communication. Changing the protocol from TCP to UDP lowered bandwidth requirements and allowed mission success. Mission success did not require optimizing one system, it required understanding the whole system.

The following are the conclusions:

- This paper was not about prediction, but rather organizing the data in order to apply analytics.
- In practice, he is proposing performance and strategy framework with existing models to reduce uncertainty and risk in using unvalidated components. The most valuable result of his work was quantifying system performance and basic information unit scales, but it is not finished yet.

3.8 Anomaly Detection of Network Traffic Based on Opaque Data

Authors: Michael Delucia, Constantin Serban, Angello Sapello, Abhrajit Ghosh, and Ritu Chadha

This paper discussed the use of ML techniques to identify malicious traffic patterns in much larger sets of benign traffic. They introduce a technique called Learning Using Privileged Information (LUPI), where they incorporate features from the individual hosts on a network into the training phase of a learning-based network anomaly detector. This additional information improves the performance (accuracy) of the detector without affecting the runtime. They demonstrate the technique on an enterprise network, where additional (privileged) information about the operating system of each host is integrated in the training phase of the ML algorithm, and then network anomalies are detected in the network with a high degree of accuracy. The method could potentially be extended to detect numerous other cyberphenomena, which might otherwise be indistinguishable from normal network background noise.

The paper demonstrates one example of using ML to try to detect network anomalies and attacks, which is the first step in predicting new attacks and then relying on analytics to determine an appropriate response. The general method might also be useful for incorporating new information about known exploits/unknown attacks into defensive or PA systems.

The following are the conclusions:

- This paper is one example of using ML to detect network anomalies and attacks.
- Improved the accuracy of ML models via privileged features available only during training.
- Capable of detecting advanced persistent threat (APT)-type stealthy malicious behavior (Doman Name System [DNS] caching example with different host operating systems).
- The example provided had excellent results, but would also have several limitations in the real world (e.g., it would not work on traffic that goes through network address translation [NAT] or over Hypertext Transfer Protocol Secure [HTTPS]).
- It is important to understand the normality of the network to focus in on anomalies because of the many events occurring.

3.9 Deep Learning Applications for Cyber Defence and Cognitive Science within the European Defence Agency (EDA) Cyber Strategic Research Agenda (SRA)

Authors: Salvador Llopis and Ignacio Montiel

This presentation was the result of an EDA-sponsored “Deep Learning Study in European Defence”. This evaluation of the current state of the art of deep learning approaches sought to 1) define a mathematical baseline that could be used for assessing performance of deep learning models, 2) analyze the use of deep learning techniques to improve automatic target recognition in radar images, 3) examine the applicability of deep learning to other defense domains for example cyber defense, and 4) provide roadmaps for deep learning implementation in the studied defense domains (radar and cyber).

The report highlighted a review of algorithms and architectures (auto-encoders, deep Boltzmann machines [DBMs], recurrent neural networks, etc.) and also reviewed some deep learning software frameworks (Caffe, Tensor Flow, Theano, etc.). There was also identification of some commercial applications (notably the "GAFA" group of Google, Apple, Facebook, and Amazon) in computer vision, natural language processing (NLP), vehicle autonomy, and healthcare. Also cited were defense applications such as object detection and tracking (optic and synthetic aperture radar [SAR] images), cyber defense, situation awareness, and detection of specific behaviors, human pose classification, speech processing, opinion mining in social networks, and improvement of autonomy of military mobile vectors. There was a detailed review of a use-case on encrypted traffic classification that will be performed over the next few years (into 2023). Some results in testing to date indicated that ML had higher score, faster training time, but longer testing time whereas deep learning had a slightly lower score, much longer training time, and faster testing time.

The second part of the talk concerned cognitive science within the EDA Cyber Strategic Research Agenda. EDA manages research and technology in 14 technology domains to develop knowledge and technologies needed for future defense capabilities. The Strategic Research and Innovation Agenda (SRIA) provides introduction to each technology domain with further detail provided in the *European Cybersecurity Research and Innovation Agenda (SRIA) for a Contractual Public Private Partnership (cPPP)* document found at http://www.ecs-org.eu/documents/ecs-cppp-sria.pdf.

The following are the conclusions:

- The authors provided a good overview of the research into deep learning within EDA. Experiments and projects that are scheduled should be monitored with respects to their outcomes.
- Of keen interest are the five research areas within the cyber situational awareness research program: dynamic risk management, decision support, CIS infrastructure discovery, cyber real-time sensor interface, and threat management.

3.10 Shaping Cyberspace: A Predictive Analysis of Adversarial of Adversarial Cyber Capabilities

Authors: Juha Kukkola, Juha-Pekka Nikkarila, and Mari Ristolainen

The authors try to estimate the implication of the Russian network on the World Wide Web as Russia is aiming to reach capability of closing its national networks, and consequently, achieving digital sovereignty. The research is based on legal procedures and official state documents (e.g., doctrines). Also publications of Russian military strategic influencers including Russian academia are considered. According to the analyzed documents, it is evident that "the Russian segment of Internet" has to be nationally controlled, independent, self-sufficient, protected from outside interference, and under sovereign jurisdictions. Effectively, Russia seeks to achieve capability to separate its national networks from the Internet by 2020 when necessary. Russia's objective is to control both its national and the global cyber domain with its own and peculiar concepts. For example "information counter struggle" (informatsionnoe protivoborstvo) is not limited solely to wartime and is different concept as its Western translation "information war". The closing process will potentially create an asymmetric situation. Essentially, when considering the traditional elements of combat power (i.e., fire power, maneuverability, and protection), it is rapidly seen that a closed-network nation may be able to achieve higher relational capabilities over its adversaries.

The following are the conclusions:

- According to the authors, Russia is manipulating cyberspace in order to achieve a decisive military advantage over its potential adversaries.
- Russia is currently manipulating the cyber domain through identified four lines of effort that are "propagating digital sovereignty, conceptual control, preparation of the cyber domain, and exploiting open society" of which the authors suggest the open society to respond by "promoting openness, conducting conceptual changes, technology improvement, and resource reorganization", respectively.
- The closing process may diminish the problem of attribution for Russia,
- The following are challenges to an active cyber defense (ACD):
    - The formation of asymmetric frontlines and shifting the freedom of action accordingly
    - An ability to control escalation by forcing an opponent to react in certain way by denying freedom of action or counterattacking

    - Reaches escalation dominance over its potential adversaries.
    - Most important question is whether or not Russia is able to find allies for this closing process.
- Although the research is novel and interesting, planning is not prediction. Planning is an analytical approach, whereas prediction is a mathematical approach. The paper is more like a conceptual paper giving an insight of possible future events.
- The group estimates that due to the closing process Russia may seek the ability to project cyber power, and then withdraw back to its own segment of Internet.
- The study is conducted mainly via a literature review method and only a qualitative method.
- The committee argues that in future studies quantitative methods are required in order to conduct better estimations.
- The problem is how to mathematically prove this model without having real data or numbers.
- Maybe applying the Delphi method would improve the analysis.
- The group suggests the NATO STO organization initiates studies in the following research areas:
    - Possible technology solutions (and their vulnerabilities) of Russia's network closing process
    - Situation awareness related to the closing process (will there be followers?)
    - Closing process influences via international legislation (e.g., the problem of attribution)

3.11 IDS Alert Prioritization through Supervised Learning

Authors: Greg Shearer, Nandi Leslie, Paul Ritchey, Tracy Braun, and Frederica Nelson

This paper presented an ML framework to assist network security analysts by automating and prioritizing alerts generated for a monitored network. The goal of their system was system to improve human analyst efficiency by prioritizing alerts and decreasing false positive reports. They used data and alerts from an intrusion detection system (IDS) that monitored an enterprise network for one month to train an ML algorithm. Then they demonstrated that their algorithm could correctly prioritize and accurately predict security incident reports over a subsequent four-month period. They showed a 99% reduction in false positives with a less than 10% reduction to true positives. The paper also notes some of the tradeoffs of accuracy versus precision when tuning the ML algorithms.

The paper demonstrated an example of using ML for threat prioritization. This type of algorithm could potentially be extended to report shifting frequencies and weights of the types of attacks seen over time. Such an algorithm could then be a useful situational awareness, trend prediction, and guided (or even automated) response. The paper also demonstrates one of the current challenges with automated ML and predictive systems—what is the threshold for missing true positives versus reducing false positives?

The following are the conclusions:

- With a properly labeled training set, they were able to increase the accuracy of security incident reports.
- With a system like this in place, analysts can spend more time looking for novel attacks and following up leads.
- Anomaly- and signature-based detection input can be fused based on an analysis of past results via event logs.
- Developing a more autonomous intrusion handling system will require both knowledge, including behavioral, criticality, and impact models, as well as the ability to gain experience (i.e., learning) by leveraging past events.
- The committee felt that if one detects trends in the strategic capabilities of an adversary, then this should also be an input to a higher level, an overall strategic threat intelligence and prediction system, with respect to this adversary's capability development/improvement.

3.12 FAST-D: Malware and Intrusion Detection for Mobile Ad Hoc Networks (MANETs)

Authors: Kenneth Yu and Nandi Leslie

The authors presented a hybrid signature- and anomaly-based IDS model called Fast Alert Signature-based Training and Detection (FAST-D). FAST-D characterizes packets in terms of n-grams and utilizes the space-efficient Bloom filter for classification. FAST-D was also designed to be "lightweight" for mobile devices, and requires less space, memory, and central processing unit (CPU) use than Snort. Experimental results showed FAST-D performed well in comparison to Snort. The FAST-D technique was also able to leverage previously known vulnerabilities to detect both known and unknown malicious activities. The tradeoff for this performance was a slightly higher rate of false positives. The FAST-D model seemed to be an effective, lightweight, and adaptable IDS for tactical and mobile devices.

The FAST-D model itself is not a PA model. It appeared to be an efficient IDS for mobile devices. However, this paper was one of the few that was found that was able to detect unknown attacks using known vulnerabilities/signatures. The n-gram and Bloom filtering was able to identify new variants of attacks without having a specific signature. This type of learning, adapting, and predicting will be necessary to implement more robust PA capabilities.

The following were the conclusions:

- FAST-D performed faster and used less memory than Snort. The tradeoff for this performance was a slightly higher rate of false positives.
- The n-gram and Bloom filtering was able to identify new variants of attacks without having a specific signature.
- Alert prioritization is needed to maintain human-supervised detection capability with lower analyst resources/increasing challenges.
- Instead of trying to establish priorities beforehand, let priorities evolve naturally. Organically growing priority focus areas has produced the most reliable threat identification.
- It includes investigation of labeling errors, reassessing false positives/false negatives for relabeling. This process will continue to drive up the accuracy of the threat data results.

4. Breakout Group and Discussions

As part of the program, the Specialist Meeting scheduled three breakout group sessions to pose discussion questions and further evaluate topics of interest. The three groups and their tasking were the following:

- Examine the establishment of a nation closing and controlling its “Internet borders” (RUNet) and how this alters current threat prediction and cyber defense. This group examined those impacts and how that would impact (manifest itself) in terms of the future of cyber defense (Group 1).

- Discuss types of predictions and how different predictive approaches would alter algorithmic implementations and outcomes (Group 2).
- Discuss how modeling and simulation could accelerate and better prepare/evaluate predictive approaches (Group 3).

Each group contained both presentation authors, Specialist Meeting attendees, and IST panel members. Each group discussion began with specific topic questions and higher-level questions for all groups. These additional questions were the following:

- What other areas of research would help and accelerate a predictive capability?
- What are potential experiments and way ahead?
- Are there any potential legal issues with respect to prediction?

The following subsections describe the discussions and conclusions of each breakout group.

4.1 Group 1: Closed and Controlled Internet Borders (RUNet) and How This Alters Prediction

The original set of questions specific to group 1 were designed to determine the predictability of cyber threats and attack vectors in the context of an environment that contains closed and controlled internet borders:

- Is the attribution problem changing after closing and controlling Internet borders (i.e., some nation(s) is/are resolving attribution problem in its/their systems)?
- What would be its impact at technical/tactical/operational/strategic levels?
- Should military planning processes be revised considering closing and controlling Internet borders after 2020?
- What is the willingness of different nations to follow the announced Russia intention?
- Would this level of control make predictions easier or harder?

The breakout group developed 11 discussion areas listed, along with the key discussion points:

1) Considering the impacts of closing and controlling Internet borders, what are the immediate outcomes of the formation of such a network?

- What trends does this set in terms of the formation of digital sovereignty(s)?
- In the specific case of Russia, who has announced intentions to create a closed and controlled Internet border (denoted RUNet), how may this impact the problem of attribution for cyber activities from closed borders?
- Increasing the complications and challenges associated with ACD from outside closed and controlled borders.
- With a lack of "geographical boundaries" in cyberspace, how will/could this action alter the formation of asymmetric frontlines and the shift in the cyberspace freedom of action both insides and outside closed and controlled borders?
- An ability to control threat escalation outside closed and controlled borders by influencing the opponent decision-making processes in certain ways, effectively denying freedom of action or counterattacking.

2) Given a situation of closed borders, what are the immediate actions to be conducted?
    - With some news reports indicating intentions of closing borders, some level of further research should be conducted on the validity and plausibility of network closing processes *and* the real/fake news reports that this is happening.
    - To better understand the operational and environmental impacts of closing borders, the breakout group discussed the construction of closed national network models.
    - As a corollary to developing models, the group also recommends the construction and testing of closed national network cyber scenarios.

3) How is the attribution problem changing as a result of closing and/or controlling Internet borders (i.e., are some nation(s) resolving the attribution problem within their systems)?
    - Issues of attribution involve identifying IP addresses and then tying actions from those addresses to governments are an already known highly complex problem. Closing the network reduces the amount of available data that can be accessed outside the closed network, although the concentration of data now comes from specific and

identifiable areas. This implies a need for reconnaissance and surveillance from inside a closed network, an area that requires further study. The answer to the question depends on how the closed national network is constructed and what other measures this causes or influences (domino effect), for example, the construction of small, closed networks across the Internet with the purpose of offensive actions. The recognition and subsequent detection of a closing/controlling process itself needs to further researched.

4) What would the impacts of closed and controlled Internet borders look like in terms of technical/tactical/operational/strategic levels?

   - Again, there was discussion on the need to build a model of a closed and controlled Internet to provide a means to evaluate impacts.

   - Any solution or approach needs to be proactive. As made in an earlier point, this further underscores the need to construct various scenarios that highlight the changes in the prediction capabilities.

   - **The group discussed the observation that if the protocols are the same, the problems will be the same. Researchers need to consider the possibility that there would be new protocols developed for closed networks. These new protocols could correct security deficiencies in current and older protocols, making interacting with closed networks even more difficult.**

5) If closed and controlled Internet borders changes the attack surface significantly, what responsive changes might have to be conducted?

   - The group discussed the potential that closed and controlled Internet borders would encourage the development of different attack vectors for offensive operations (either inside or outside the closed network), including small deployed closed offensive networks across the Internet. There was also some discussion on how this might impact or alter known TTPs for insider threats.

   - One topic discussed was to take a contrary position to discuss if a closed national network could potentially improve prediction. Part of this contention was based on the notion that a closed national network owner needs to have almost complete situational awareness of their own system in order to maintain full cyber control. Along with this control, a point was raised about a potential side effect where homogeneity could create new vulnerabilities.

6) Should cyber military planning processes be revised after the implementation of closing and controlling Internet borders outlined in the RuNet 2020 report?

    - Again, before any action can be considered, solid proof of the effects is required before the processes are to be revised.

    - The group discussed the need to continue researching closed and controlled Internet borders because the intelligence gathering process will be impacted thus affecting accurate modeling and simulation of plans and cyber wargaming. In this case, a closed network could require additional capabilities to complete the planning process.

    - With respect to NATO actions, discussion centered on the need to define what is a member responsibility versus what would be a NATO responsibility.

7) Discussion then proceeded to evaluate the willingness of different nations to follow a closed and controlled Internet border solution:

    - Further emphasis was placed on establishing modeling and simulation of closed networks to develop situational awareness, understanding, and recognition of precursors to a closed and controlled Internet border.

    - This also inferred several concerns: If one's attempt at a closed and controlled Internet border succeeds, there is a potential for other occurrences. Some societies would accept the constraints of a closed and controlled Internet border; others may not due to the tradeoff of security versus freedom.

    - There was also voiced a concern that countries might be pressured or blackmailed into closing network borders.

    - The economic benefit from closing network borders is not accurately known or researched.

8) Would this level of control make predictions of adversarial behavior easier or harder?

    - **Most likely it would make predictions easier, but the closed network nation would need to form additional measures to address the problem. The owners of a closed network have better**

visibility into ingress/egress points of that network to predict offensive operations targeted against the closed network.

9) What other areas of research on closed and controlled Internet borders would help?

   - Are we missing the target looking at prediction? PA may be viable for zero-day technical exploits, but not as viable for TTP zero-day exploits.

   - How do we generate valid training data? This could indicate a paradigm shift in the types and format of realistic testing data and there will not be a body of knowledge on what a new attack vector coming from a closed and controlled Internet border would "look like". In addition, the command and control structures used with or against a closed and controlled Internet border network are unknown.

10) What are potential experiments and the way ahead?

   - Create a model of a closed national network and test its features and implications in different scenarios.

11) Are there any potential legal issues with respect to prediction in closed and controlled Internet borders?

   - While not having been evaluated, the group felt the answer is most likely yes. However, there needs to also be legal assessment of closed and controlled Internet borders.

The group concluded that this is an important issue to be solved: Is it possible to actually reach escalation dominance via closed national network over nations within open society?

- Discussion and Recommendations:

   - Open-network nations need to collectively understand and have potential responses to nations who will close their network borders, otherwise, the open-network society may lose the ability to influence and fully understand the cyber domain. Developing a better understanding of closed networks, their characteristics and footprints, precursors to network closing events, and impacts to current cyber defense and intelligence gathering will be required. The recommendation is to develop models of closed networks to facilitate this research.

- Enhance the intelligence gathering on nations openly discussing closed networks. This includes more authentic published documents (doctrines, state strategies and programs), legislation (bill drafts, other documents), along with works by leading scientists and researchers.
- **Internet fragmentation is the de facto ongoing process and RuNet is predicted to be in operational use as per 2020.**
- The NATO STO organization is initiating studies on the following research areas:
    - Possible technology solutions (and their vulnerabilities) of the closed national networks closing process
    - Situational awareness related to the closing process (e.g., will there be followers?)
    - Closing process influences via international legislation (e.g., the problem of attribution)
    - Closing process influences on operational capabilities

4.2 Group 2: Types of Prediction Breakout Group

The second breakout group during the workshop discussed types of prediction with respect to PA. Questions posed to the breakout group included the following:

- What are the implementation/operational issues of having an operationally feasible (and over 80% accuracy) predictive set of analytics?
- How will this change cyber defense?

The discussion pointed out several issues:

- There appeared agreement that ML is a viable methodology to attain reasonable levels of prediction of the (known, unknown) capability threat.
- The threat can adapt and change vectors faster than the algorithms or learning can react. This implies the shelf life of the processes are in question in terms of changing threat vectors adversaries are developing.
- There is a question as to what 80% accuracy means. Is it in terms of predicting 80% of the events correctly? Does it mean 80% of tipping/queuing of analysts correctly? The accuracy of prediction could be interpreted in several ways.

- The Specialist Meeting presented information that algorithms and processes to identify and possibly predict threats in the (known, unknown) region of the DSB definition will be multi-method ensembles as opposed to single-method algorithms. This goes against conventional thinking of single-method algorithms to identify all threats in cyberspace.

- There are temporal issues concerned with prediction. The ability to perform predictions must be in real time or near real time. If too much time passes, the prediction could be overcome by other cyber events.

- The discussion pointed out that there would need to be advances in processing speed and power along with algorithmic refinement to better address the temporal implications.

- If the adversary understood and could manipulate the predictive approaches, they would be able to generate any situational awareness they desired. Generation of false positives impacting ML could disrupt the thresholds for the decision trees and create meaningless courses of action.

- Considering the previous point, resiliency approaches to the integrity of the prediction process need to be developed alongside the prediction approaches themselves.

- The concept of separating algorithms that are resource inefficient (use more power, memory, time, etc.) from more resource-efficient algorithms could be used in combination to produce results balancing accuracy and response time. This is an area that would benefit from more research.

- The more capable the threat (in terms of level within the DSB model), the less deterministic their behavior, making them more difficult to identify and/or predict.

Several additional questions were posed to the group. The questions and resulting discussions are summarized.

Question: Extending the concepts to the threats that invent new vulnerabilities with new exploits (unknown, unknown), is there a way ahead for analytics?

- While not discussed in detail, the threat with (unknown, unknown) capabilities may be an area better suited for artificial intelligence and ML algorithmic approaches. Developing data sets for the (unknown, unknown) threat will be complex and challenging impacting the accuracy of learning algorithms due to smaller and incomplete data sets.

- Based on the discussion about the (unknown, unknown) threat identification, we would expect an increase of cyber defense resource consumption above the (known, unknown) threat.
- Some discussion mentioned human analysts in the (unknown, unknown) loop and that a benefit may be gained from better visualizations of the data being analyzed.

Question: What other areas of research would help and accelerate a predictive capability?

- The development of a matrix of implementation approaches (i.e., Bayesian, Monte Carlo, supervised learning, etc.) against what kinds of problems they best solve to guide algorithmic implementation approaches (i.e., which implementation methods are better suited against specific prediction environments).
- It would be useful to research what are the mechanisms that define the boundaries of threat capabilities. What are the characteristics or observed abilities that could define a threat against the three threat capability levels as defined by the DSB report *Resilient Military Systems and the Advanced Cyber Threat*.
- It appeared during the discussion that using attack graphs as part of the prediction problem might help in reducing resource usage and more importantly increase prediction accuracy and reduce resultant deviation (predictive stability).

Question: What are potential experiments and the way ahead?

- More detailed investigation into the structure of analytics from the big four companies: Google, Apple, Facebook, and Amazon (GAFA, per European Union and presentation from Salvador Llopis; some add Netflix to create FAANG) to see if any of those processes are useful to the cyber prediction problem.

The discussion led the group to attempt to define and build up a theoretical model for a way measure and compare the effectiveness of different predictive techniques in an operational environment, incorporating inputs, outputs, risk, and response. A rough outline of the model included the following:

- Inputs: Accuracy of prediction (true positive vs. false positive) and mission goals, including confidentiality, integrity, and availability.

- Outputs: Decisions (courses of action), recommendations, strategic innovations, and disaster recovery.
- Risk factors: Human power, laws and regulations, operational risk, facilities, and technology.
- Response: Cost, return on investment, impact, likelihood of success, moral implications, and political impact.

Connecting all these components together with a feedback loop in the model could be used to refine prediction capabilities.

4.3 Group 3: Modeling and Simulation Breakout Group

The third breakout group during the workshop discussed modeling and simulation with respect to PA. Questions posed to the breakout group included the following:

- What are the modeling and simulation requirements to adequately testing/developing predictive systems?
- If PA were used to identify attacks, could that be constantly run to not only identify potential attacks, but to generate the correct patches (rendering things the known-known signature based situation)? How would that be done? How does that change cyber defense as we know it?
- Should this be done in real time?
- Additional information:
    - What other areas of research would help and accelerate a predictive capability?
    - What are potential experiments and the way ahead?
    - Are there any potential legal issues with respect to prediction?

Several of the members of the breakout group had attended the recent NATO Modelling and Simulation Workshop in July 2017. This workshop was held in conjunction with IST-156/MSG-151 by Dr Ritu Chadha (IST) and Mr Jack Bramhill (MSG). The following comments were made during the breakout session:

- Traditionally, modeling and simulation concentrated on physical effects. Modeling and simulation in the cyber domain is less mature and must consider adversarial behavior, either human or future intelligent system behavior.

- Recommendations from the NATO Modelling and Simulation Workshop (July 2017) included the following:
    - Formation of an ET to produce a "top 10" list of effects/attacks whose representation should be prioritized in future work.
    - Formation of a NATO HMF/IST/MSG workshop for common symbology, taxonomy, and standards.
    - Formation of ETs to study the applicability of current international law to cyberspace, need for new regulations, and how trends will affect the future cyberspace operating environment.
- Considering the previous three recommendations and applying them to effective modeling and simulation for PA, the following insights were provided:
    - PA would need a common terminology. Predictions should have a certain percentage, with certain modality, with certain truth value, and so on. Predictions also need to have a temporal element and should have level of accuracy. For example, the "seismologist problem" occurs with earthquake predictions. How useful is a prediction for a major earthquake in the next week (vs. in the next month or next year)?
    - Another requirement is the need to measure and define performance. How should a successful prediction be defined?
- When talking about predictions, anything is possible (within certain universal limitations). Predictions in cyberspace are not limited to certain physical or temporal constraints. Cyberspace does not have traditional physical constraints. Because (almost) anything is possible in the future in cyberspace, the space of possible (if unlikely) outcomes is extremely large. Therefore, this space is difficult to model and simulate.
- Sharing of data among NATO partners is difficult. Different countries have different concerns and different problems sharing data. Thus, determining ground truth of data sets is tough. This is a place where modeling and simulation can be useful. It can be used to generate sharable data with known statistics.
- One requires an accurate model of a system before it can be studied and predictions generated. Current enterprise network models work well. But that is not enough for PA. Other factors must also be modeled (e.g., threats, vulnerabilities, and adversaries). However, once a model is too complex,

the system is just being rebuilt/replicated. Modeling and simulation should abstract and simplify somewhat.

- There are currently good system models for physical systems, such as tanks. But there are not good models for systems in cyberspace. Tanks have well-understood physical constraints. Cyberspace does have some physical constraints (e.g., bandwidth). However, they are not the traditional physical constraints. Therefore, they are not well understood, and it will require new research to study these structures and restrictions.
- Combining risks (e.g., from a "top 10" list of prioritized attacks of concern) and determining impact is another area where modeling and simulation can be especially helpful.
- In the short term, for PA in cyberspace, researchers need to start with a simple model, simple predictions, and build up to more complex simulations from there.

For the discussion question, "If one has successful prediction system, can one predict future patches?", the group commented as such:

- A predictive system might be able to generate certain types of patches. More likely, a predictive system could help prioritize future patching and allocation of defensive resources. The system would need measurement of applied security controls, so the system can learn if that was an effective response.
- The following counter example to the question was also given. If one has a successful PA algorithm running constantly, that can be obtained and used by an adversary. An adversary can take that algorithm and manipulate inputs to exploit against it. So, no system can make perfect predictions.
- A predictive system must also be able to detect and measure certain levels of adversarial deception and react accordingly.

For the discussion question, "Can a successful prediction system be run in real time?", the group commented as follows:

- If one is doing prediction, one needs some sort of modeling and simulation to evaluate the impact of the maneuver and the effectiveness of the prediction. Computer horsepower is still a problem, but the models and algorithms also need work. Researchers in PA should leverage the computer gaming industry more. They model complex interactions with humans and complex systems on simple gaming systems. Can we leverage them more?

5. Conclusions and Findings

This workshop identified several areas that are researching prediction both within and outside the cyber domain. While some work has been done, not essentially enough in the opinion of the committee, much work still needs to be done in both research and implementation. Our results from this Specialist Meeting are identified below and organized into five areas: key results and findings as identified by the committee, some general observations on the practice of prediction, and then some recommendations for the cyber modeling, cyber analytics/algorithms, and cyber prediction communities:

- Key results and committee findings:
  - Several papers introduced multiple algorithmic approaches, for example, one paper described a two-model approach with one checking uniformity of the model with statistically proven method (Bayesian) whereas the other is checking the autocorrelation (Monte Carlo). Our discussions both during presentations and in the breakout groups concluded that it appears no one algorithm is enough to solve problem. This appears to support the notion that correlation in cyberspace will not use single-algorithmic approaches.
  - The committee felt that the specific edge cases of 0% certainty an event will not happen and 100% certainty that an event will happen might be unattainable. This is primarily due to the possibility that events can be independent variables in the computations.
  - The committee also concluded that the known vulnerability/known exploit is a solvable problem and has been solved, but not necessarily implemented through automation. It is detection as opposed to prediction, making prediction in this case trivial.
  - The committee felt that if one detects trends in the strategic capabilities of an adversary, then this should also be an input to a higher-level, overall strategic threat intelligence and prediction system with respect to this adversary's capability development/ improvement and possible new or altered cyber TTPs.
  - The committee felt that the structure of STIX lends itself to more efficient communications across all entities working the cyber event prediction problem. STIX already is structured to contain important information and was formed to help security practitioners "to better

understand what computer-based attacks they are most likely to see and anticipate and/or respond to those attacks faster and more effectively ”

- **ML approaches are capable of detecting APT-type stealthy malicious behavior. For example, Zhao et al. (2015) used DNS traffic and traffic analysis ML to detect APTs.**

- General observations from the Specialist Meeting:
  - Papers mostly addressed analytical approaches with varying degree of application to the known vulnerability, unknown exploit problem.
  - Identifying and understanding a baseline security posture is important to understand the normal state of the network as the initiator to focus on anomalies that deviate from that normal state.
  - There is some important research being performed, particularly within EDA, DARPA, and other national research agencies. This work should be monitored and outcomes shared. Key cognitive application areas being investigated may include artificial intelligence for cyber operations, ML for cyber operations; deep learning (neural networks) for cyber operations, human factors for cyber defense, and algorithms design and engineering.
  - Instituting RuNet approaches can adversely affect the ability to do prediction, event correlation, and attribution.
  - Most all discussions mentioned the lack of valid training data or at least sets of training data where the validity and provenance were certain.
- Recommendations to the cyber modeling community:
  - Different SSMs require different number of samples for operating at the same level of accuracy (even the same SSM at different states). In addition, recent advances in MIS and AIS allow one to use few samples and still have a great performance.
  - Developing a more autonomous intrusion handling system will require both knowledge, including behavioral, criticality, and impact models, as well as the ability to gain experience (i.e., learning) by leveraging past events.

    - Models of a closed (national-level) network and construct representative cyberattack scenarios. By doing that, we may be able to extract characteristics of closed-network spaces.
    - Modeling and simulation of potential predictions could provide insight into affects and effects of acting on a particular prediction.
- Results for the cyber analytics and algorithm community:
    - Given the approach from the Bowman paper (that analyzes relationship entities to identify potential members of a threat group), the committee agreed that this approach for this use case is useful in clarifying relationships. Potential application in cyber domain is not so much prediction, but rather given a set of cyber events (the "messages" from this paper) what could be hierarchy or the relationship across those events.
    - Developing attack graphs around known vulnerabilities could generate all or most all of the possible attack paths. This approach may be able to reduce the prediction problem (for the known, unknown case only) to a more deterministic approach that concentrates on likelihood of a graph event occurring.
    - Anomaly- and signature-based detection inputs can be combined based on an analysis of past results of event logs.
    - Some discussion pointed out that if an adversary compromised the PA, that adversary could manipulate inputs thereby exploiting the algorithm and negating its effectiveness. In fact, they could manipulate the inputs to maneuver the cyber defender into a more vulnerable position (cyber deception).
- Results addressing cyber prediction:
    - Although some of the research is novel and interesting, planning is not prediction. Planning is analytical and partial mathematical approach whereas prediction results are better served via a mathematical approach.
    - Using the known vulnerabilities as a mechanism to produce attack graphs identifying potential exploits can reduce the space of uncertainty in predictions.
    - When talking about predictions, anything is possible (within certain universal limitations). Predictions in cyberspace are not limited to

certain physical or temporal constraints. Cyberspace does not have traditional physical constraints. Because (almost) anything is possible in the future in cyberspace, the space of possible (if unlikely) outcomes is extremely large. Therefore, this space is difficult to model and simulate.

- Discussions indicated we may not be able to predict with certainty, but we may be able to predict likelihood.

- A predictive system could be applied to other areas of cyber defense to potentially help prioritize future patching and allocation of defensive resources including identification of adversarial deception and use the PA to select potential courses of action.

6. References

Bell B, Santos E Jr, Brown SM. Making adversary decision modeling tractable with intent inference and information fusion. Adversarial Intent Inference for Predictive Battlespace Awareness. 2005;4.

Brown SM, Santos E Jr. Bell B. Knowledge acquisition for adversary course of action prediction models. In Proceedings of the AAAI 2002 Fall Symposium on Intent Inference for Users, Teams, and Adversaries; 2002.

Bushman JB, Mitchell CM, Jones PM, Rubin KS. ALLY: an operator's associate for cooperative supervisory control systems. IEEE Trans Sys Man Cybernetics. 1993;23(1):111–128.

Carberry S. Modeling the user's plans and goals. Computational Linguistics. 1988;14(3):23–37.

Chu RW, Mitchell CM, Jones PM. Using the operator function model and OFMspert as the basis for an intelligent tutoring system: towards a tutor/aid paradigm for operators of supervisory control systems. IEEE Trans Sys Man Cybernetics. 1995;25(7):1054–1075.

Geddes ND. A model for intent interpretation for multiple agents with conflicts. Proceedings of IEEE International Conference on Systems, Man and Cybernetics; 1994 Oct 2–5; San Antonio (TX). IEEE. 1994; 3:2080–2085.

Geddes N. The use of individual differences in inferring human operator intentions. AAAIC'86- Aerospace Applications of Artificial Intelligence; 1986. p. 31–41.

Goodman BA, Litman DJ. Plan recognition for intelligent interfaces. Sixth Conference on Artificial Intelligence Applications. IEEE; 1990. p. 297–303.

Gosler JR, Von Thaer L. Task force report: resilient military systems and the advanced cyber threat. Washington (DC): Department of Defense, Defense Science Board; 2013. p. 41.

Kott A, McEneaney WM, editors. Adversarial reasoning: computational approaches to reading the opponent's mind. Boca Raton (FL): CRC Press; 2006.

Lesh N, Rich C, Sidner CL. Using plan recognition in human-computer collaboration. UM99 User Modeling. Vienna: Springer; 1999. p. 23–32.

Lesh N, Martin N, Allen J. Improving big plans. AAAI/IAAI; 1998 July. p. 860–867.

Linkov I, Eisenberg DA, Plourde, K, Seager, TP, Allen J, Kott A. Resilience metrics for cyber systems. Environment Systems and Decisions. 2013;33(4): 471–476.

Rubin KS, Jones PM, Mitchell CM. OFMspert: Inference of operator intentions in supervisory control using a blackboard architecture. IEEE Trans Sys Man Cybernetics. 1988a;18(4):618–637.

Rubin RB, Perse EM, Barbato CA. Conceptualization and measurement of interpersonal communication motives. Human Communication Research. 1988b;14(4):602–628.

Santos E Jr, Zhao Q. Adversarial models for opponent intent inferencing. In: Kott A, McEneaney. Adversarial reasoning: computational approaches to reading the opponents mind. Boca Raton (FL): CRC Press; 2006 July 20. pp. 1–22.

Santos E. A cognitive architecture for adversary intent inferencing: Structure of knowledge and computation. In Enabling Technologies for Simulation Science VII; 2003 Sep. International Society for Optics and Photonics. 2003;5091:182–194.

Zhao G, Xu K, Xu L, Wu B. Detecting APT malware infections based on malicious DNS and traffic analysis. IEEE Access. 2015:3;1132–1142.

Appendix. Presentations

This appendix appears in its original form, without editorial change.

NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

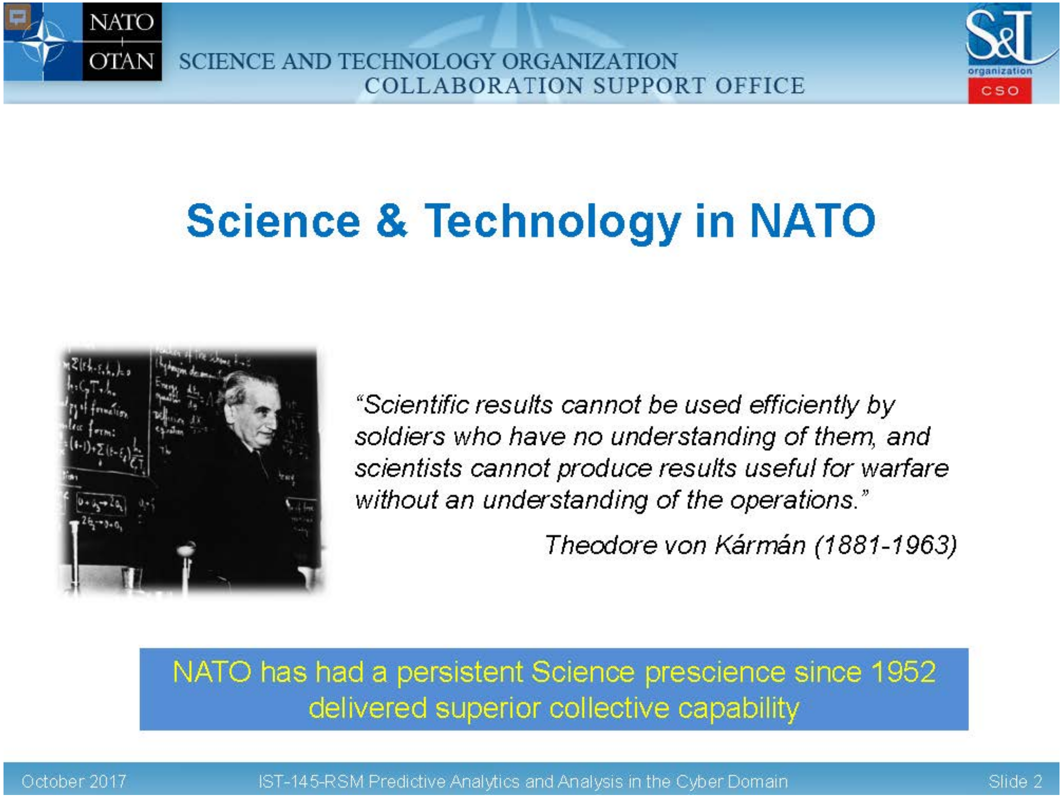
NATO Science & Technology Organization

IST Collaborative Network and the Collaborative Support Office

Dr. Michael WUNDER

Information Systems Technology Panel Chair

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 1



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

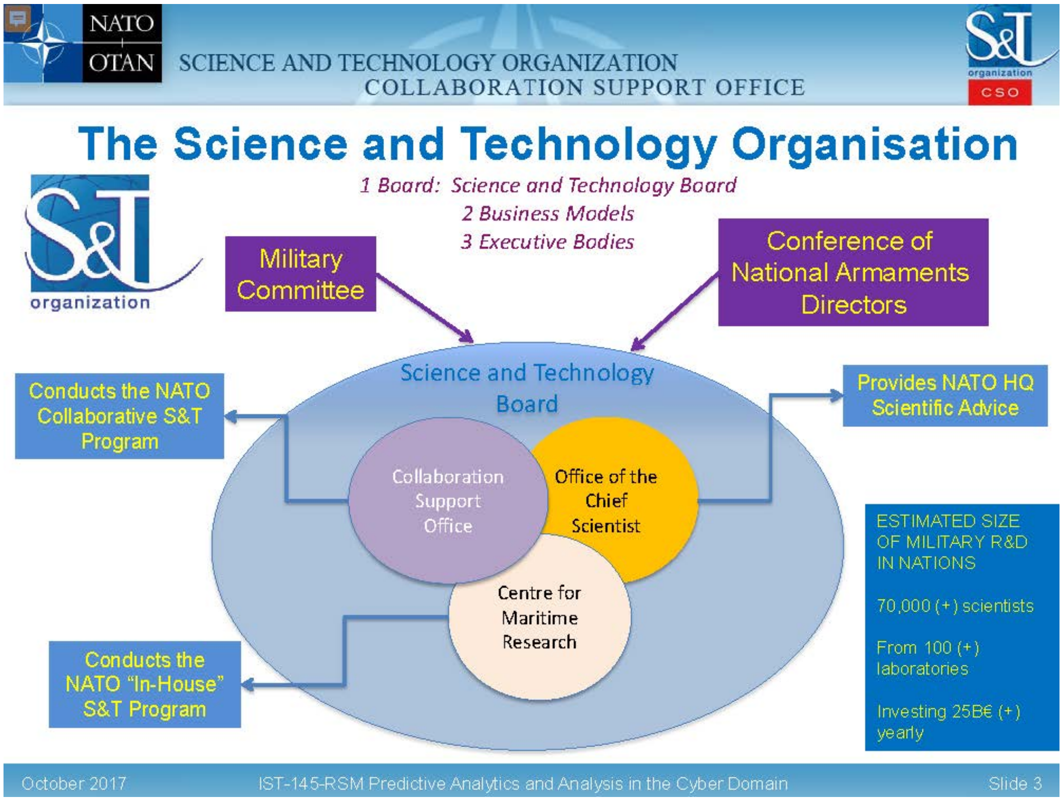
Science & Technology in NATO

"Scientific results cannot be used efficiently by soldiers who have no understanding of them, and scientists cannot produce results useful for warfare without an understanding of the operations."

Theodore von Kármán (1881-1963)

NATO has had a persistent Science prescience since 1952 delivered superior collective capability

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 2



The Science and Technology Organisation

1 Board: Science and Technology Board
2 Business Models
3 Executive Bodies

Military Committee

Conference of National Armaments Directors

Science and Technology Board

Collaboration Support Office

Office of the Chief Scientist

Centre for Maritime Research

Conducts the NATO Collaborative S&T Program

Provides NATO HQ Scientific Advice

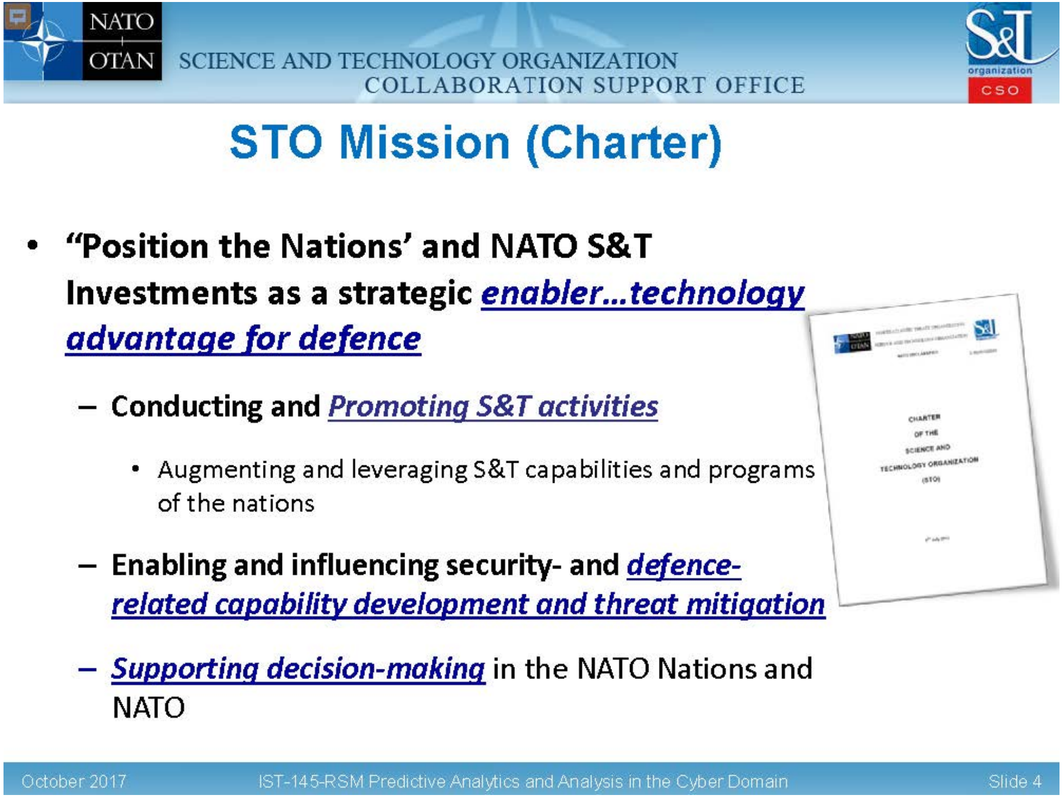
Conducts the NATO "In-House" S&T Program

ESTIMATED SIZE OF MILITARY R&D IN NATIONS

70,000 (+) scientists

From 100 (+) laboratories

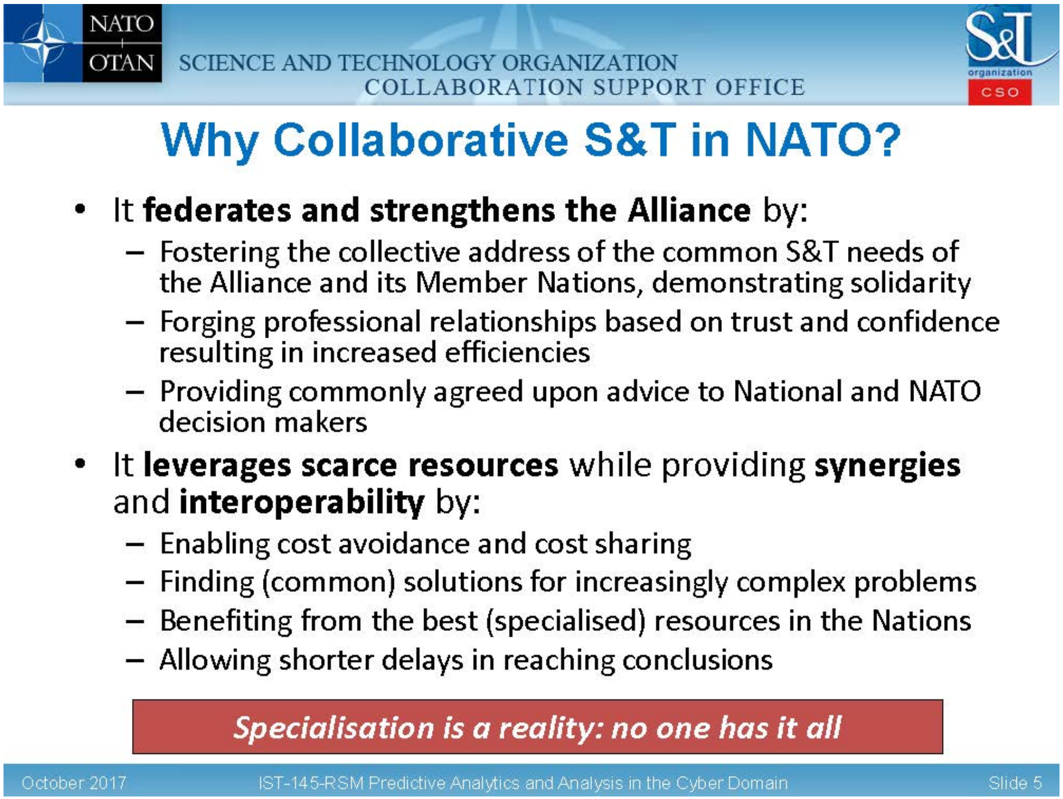
Investing 25B€ (+) yearly



STO Mission (Charter)

- **"Position the Nations' and NATO S&T Investments as a strategic *enabler…technology advantage for defence***
  - **Conducting and *Promoting S&T activities***
    - Augmenting and leveraging S&T capabilities and programs of the nations
  - **Enabling and influencing security- and *defence-related capability development and threat mitigation***
  - ***Supporting decision-making* in the NATO Nations and NATO**

CHARTER OF THE SCIENCE AND TECHNOLOGY ORGANIZATION (STO)



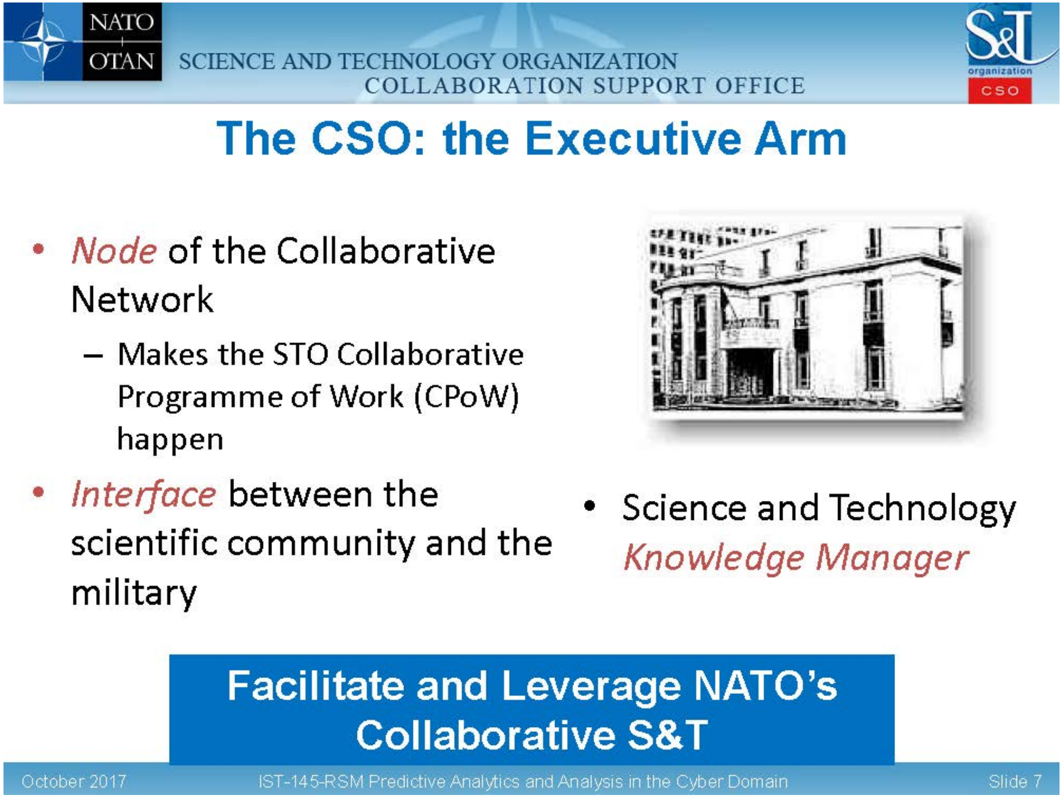
NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

Why Collaborative S&T in NATO?

- It **federates and strengthens the Alliance** by:
  - Fostering the collective address of the common S&T needs of the Alliance and its Member Nations, demonstrating solidarity
  - Forging professional relationships based on trust and confidence resulting in increased efficiencies
  - Providing commonly agreed upon advice to National and NATO decision makers
- It **leverages scarce resources** while providing **synergies** and **interoperability** by:
  - Enabling cost avoidance and cost sharing
  - Finding (common) solutions for increasingly complex problems
  - Benefiting from the best (specialised) resources in the Nations
  - Allowing shorter delays in reaching conclusions

Specialisation is a reality: no one has it all



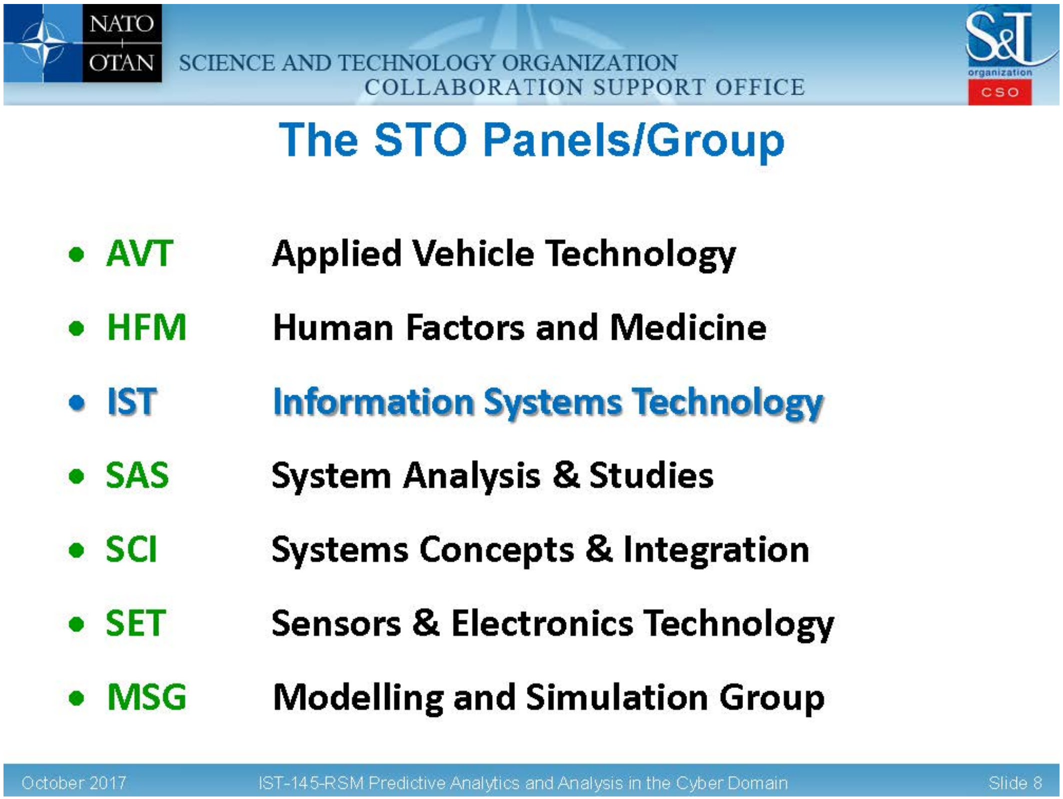
NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

The CSO: the Executive Arm

- *Node* of the Collaborative Network
  - Makes the STO Collaborative Programme of Work (CPoW) happen
- *Interface* between the scientific community and the military
- Science and Technology *Knowledge Manager*

Facilitate and Leverage NATO's Collaborative S&T

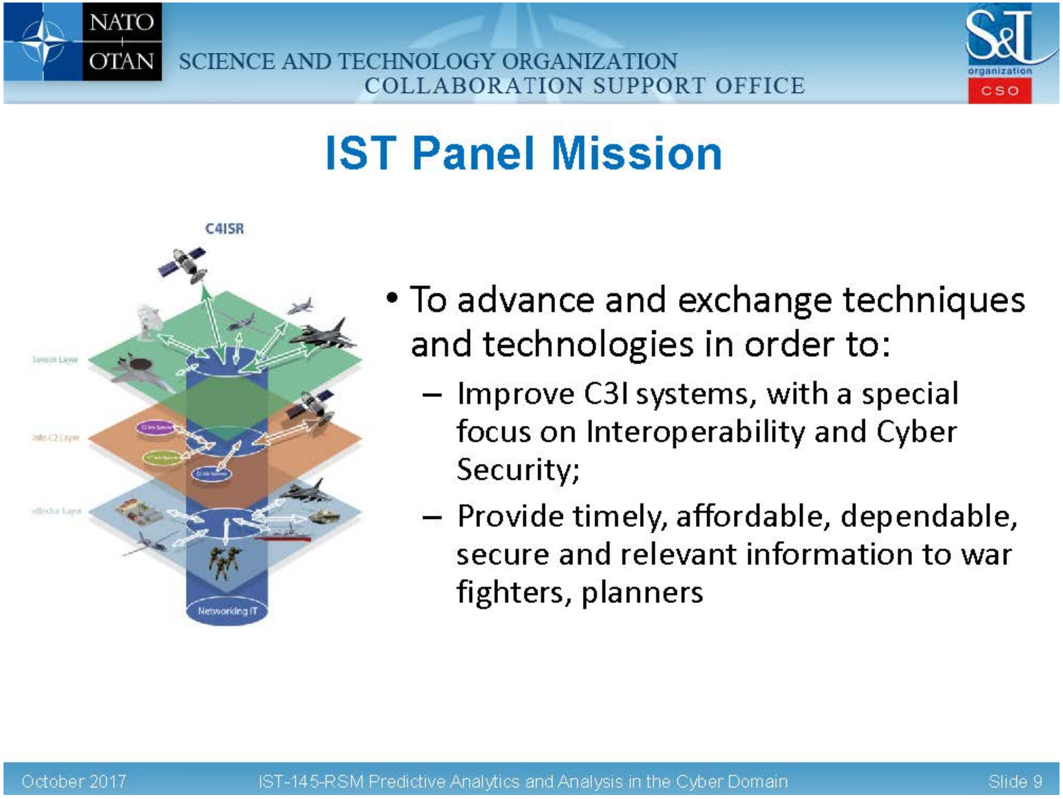


NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

The STO Panels/Group

• AVT	Applied Vehicle Technology
• HFM	Human Factors and Medicine
• IST	Information Systems Technology
• SAS	System Analysis & Studies
• SCI	Systems Concepts & Integration
• SET	Sensors & Electronics Technology
• MSG	Modelling and Simulation Group

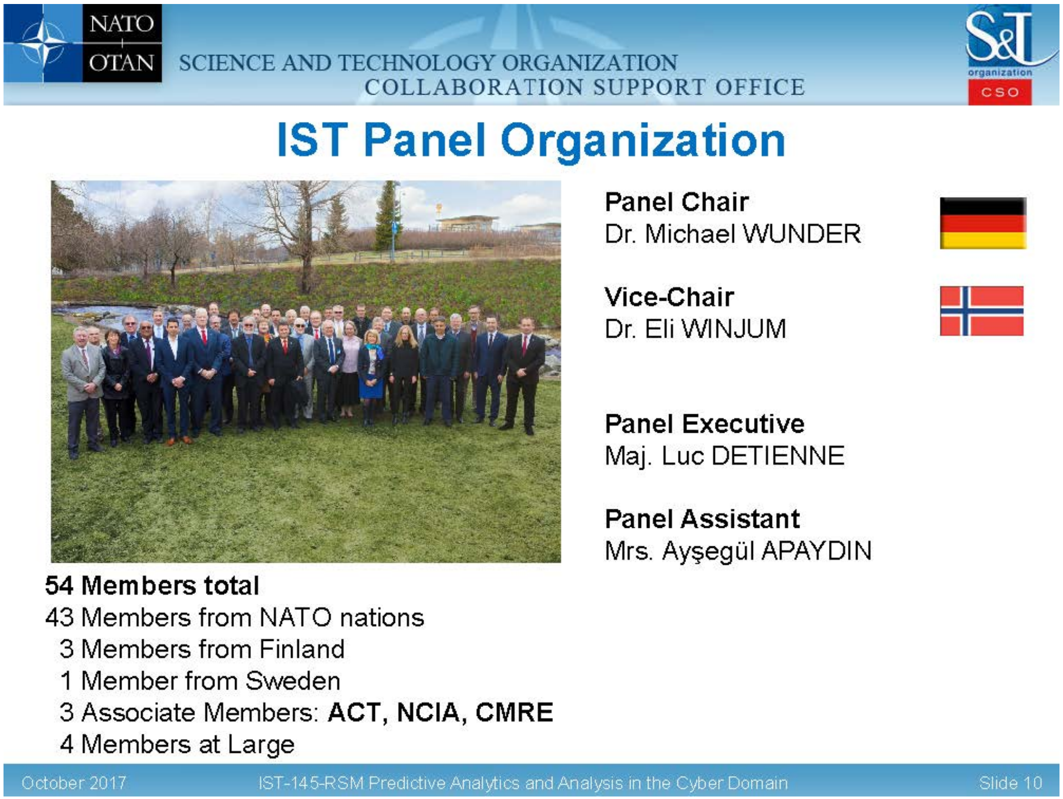

C4ISR
Sensor Layer
Info C2 Layer
Effector Layer
Networking IT

NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

IST Panel Mission

- To advance and exchange techniques and technologies in order to:
  - Improve C3I systems, with a special focus on Interoperability and Cyber Security;
  - Provide timely, affordable, dependable, secure and relevant information to war fighters, planners



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

IST Panel Organization

Panel Chair
Dr. Michael WUNDER

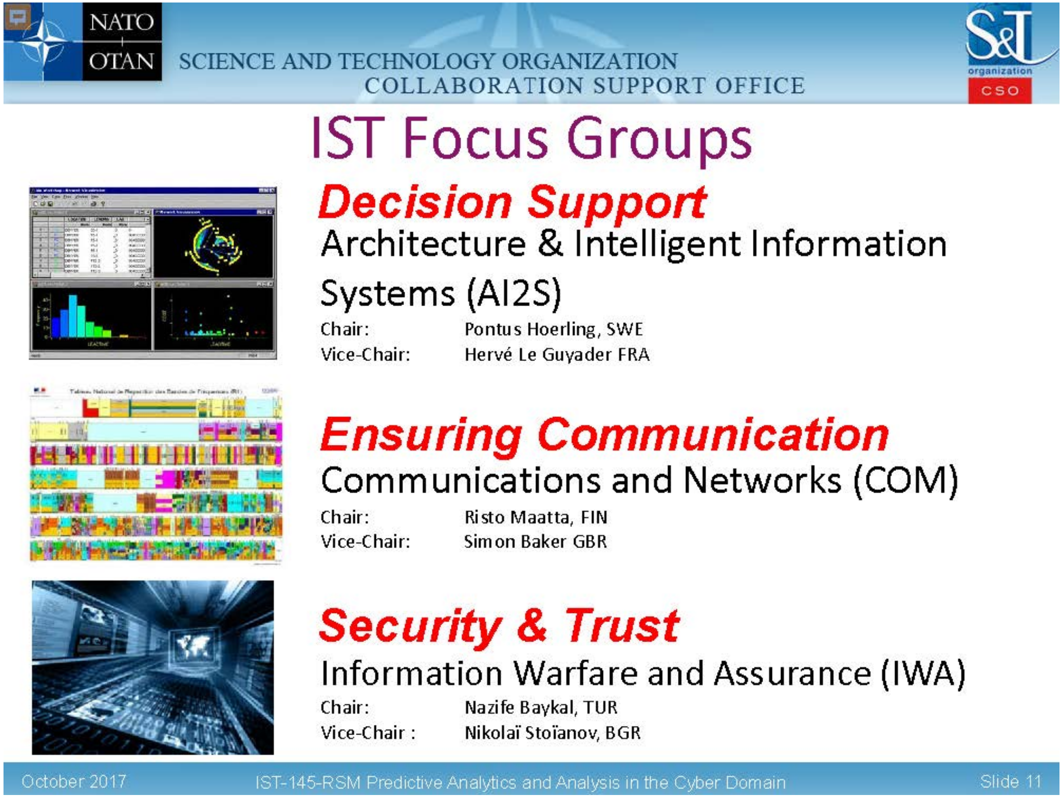
Vice-Chair
Dr. Eli WINJUM

Panel Executive
Maj. Luc DETIENNE

Panel Assistant
Mrs. Ayşegül APAYDIN

54 Members total
43 Members from NATO nations
3 Members from Finland
1 Member from Sweden
3 Associate Members: **ACT, NCIA, CMRE**
4 Members at Large

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 10



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

IST Focus Groups

Decision Support
Architecture & Intelligent Information Systems (AI2S)

Chair:	Pontus Hoerling, SWE
Vice-Chair:	Hervé Le Guyader FRA

Ensuring Communication
Communications and Networks (COM)

Chair:	Risto Maatta, FIN
Vice-Chair:	Simon Baker GBR

Security & Trust
Information Warfare and Assurance (IWA)

Chair:	Nazife Baykal, TUR
Vice-Chair :	Nikolaï Stoïanov, BGR

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 11

How to Participate

Exploratory Teams (ET) assist or advise the Panel on the technical merit or feasibility of a specific proposal for a technical activity. Exploratory Teams may also be used to help the Panel develop recommendations on future content of the Panel's technical programme.

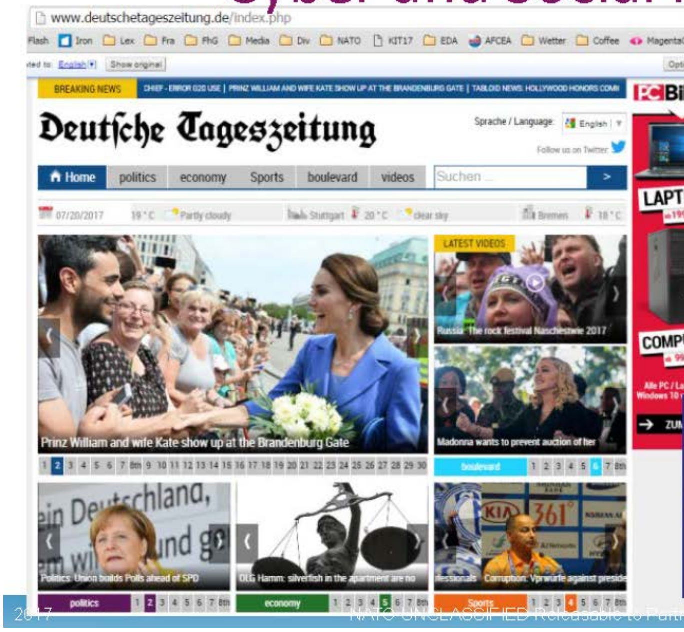
A **Research Task Group (RTG)** is chartered for a maximum of three years to solve a particular research and technology problem. The findings will be documented in a RTO publication (Technical Report or Technical Memoranda).

A **Research Symposium (RSY)** promotes the exchange of state-of-the-art knowledge among a wide audience on an important scientific or applied topic.

Other possibilities include **Lecture Series**, **Workshops**, **Specialists' Meetings**, etc..

Current and Future Focus
Recent Accomplishments
New Activities

Cyber and Social Media



CHIEF - ERROR 020 USE | PRINZ WILLIAM AND WIFE KATE SHOW UP AT THE BRANDENBURG GATE | TABLOID NEWS: HOLLYWOOD HONORS COMI

Prinz William and wife Kate show up at the Brandenburg Gate

Russia: The rock festival Naschestwie 2017

Madonna wants to prevent auction of her

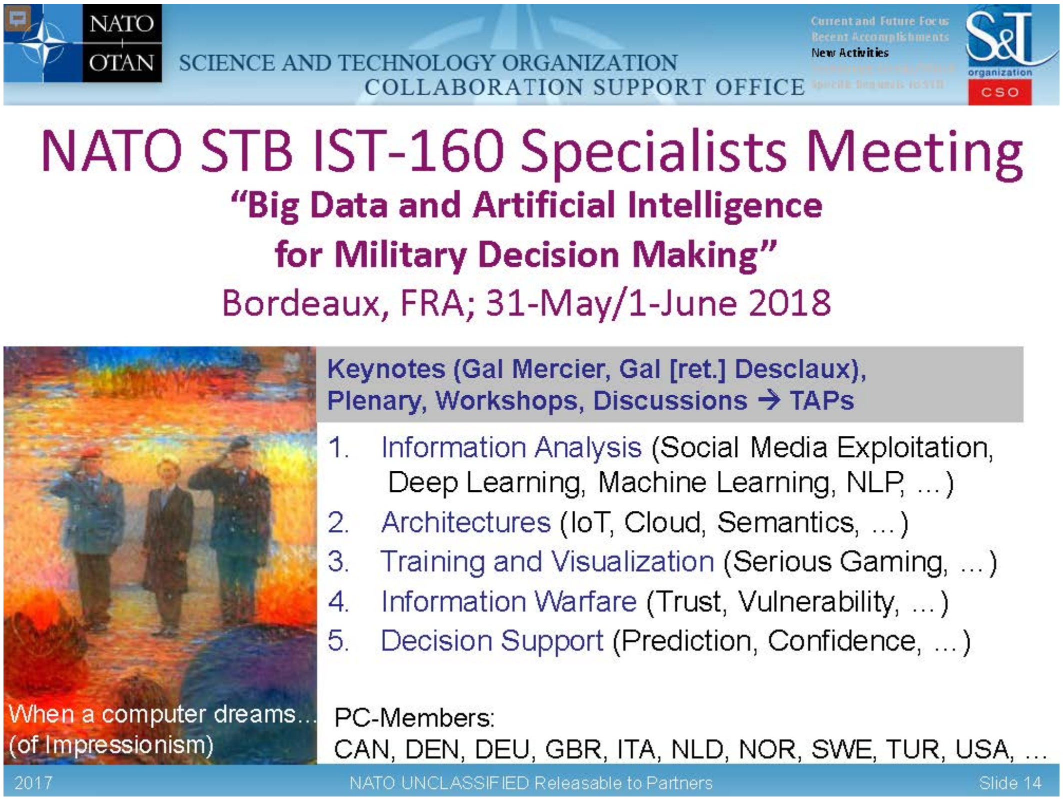
Politics: Union builds Polls ahead of SPD

OLG Hamm: silverfish in the apartment are no

Corruption: Vorwürfe against preside

- NW Analysis
- Social Media Exploitation
- Background Knowledge
- Consider Framework

→ *IST-159 RTG on Cyber Intelligence and Social Media*



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

Current and Future Focus
Recent Accomplishments
New Activities
[illegible]
Specific Requests to STB

S&T organization CSO

NATO STB IST-160 Specialists Meeting

"Big Data and Artificial Intelligence for Military Decision Making"

Bordeaux, FRA; 31-May/1-June 2018

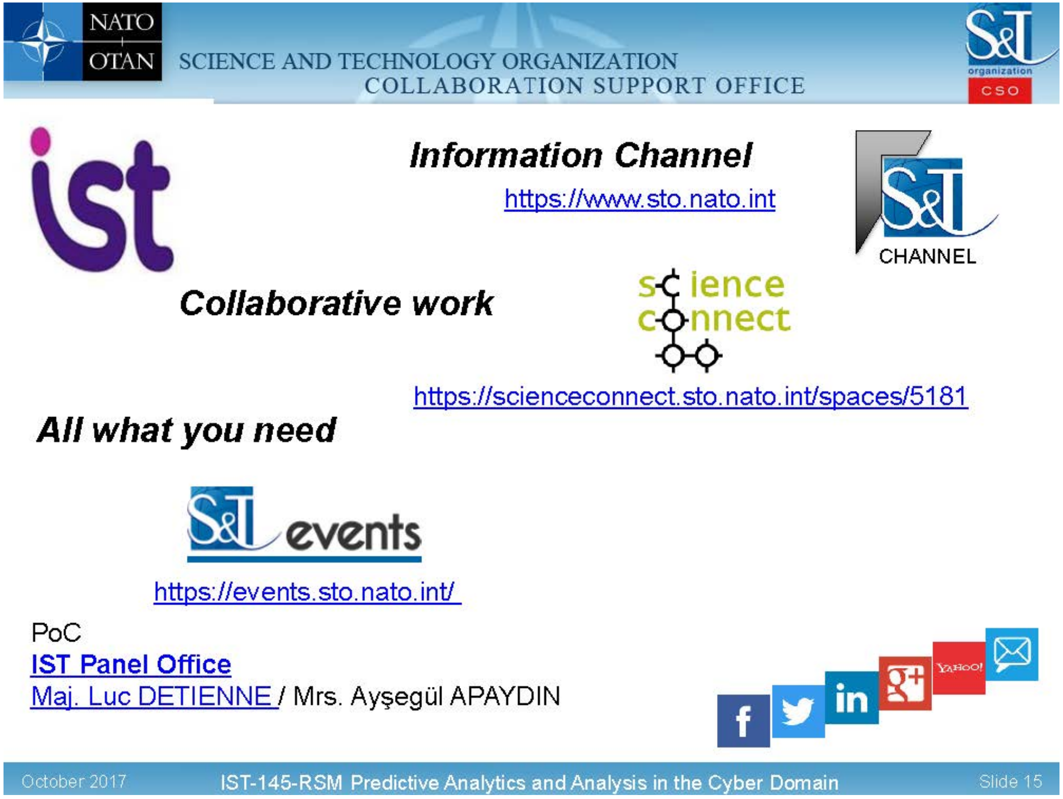
When a computer dreams… (of Impressionism)

Keynotes (Gal Mercier, Gal [ret.] Desclaux), Plenary, Workshops, Discussions → TAPs

1. Information Analysis (Social Media Exploitation, Deep Learning, Machine Learning, NLP, …)
2. Architectures (IoT, Cloud, Semantics, …)
3. Training and Visualization (Serious Gaming, …)
4. Information Warfare (Trust, Vulnerability, …)
5. Decision Support (Prediction, Confidence, …)

PC-Members:
CAN, DEN, DEU, GBR, ITA, NLD, NOR, SWE, TUR, USA, …

2017 NATO UNCLASSIFIED Releasable to Partners Slide 14



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

ist

Information Channel

https://www.sto.nato.int

S&T CHANNEL

Collaborative work

science connect

https://scienceconnect.sto.nato.int/spaces/5181

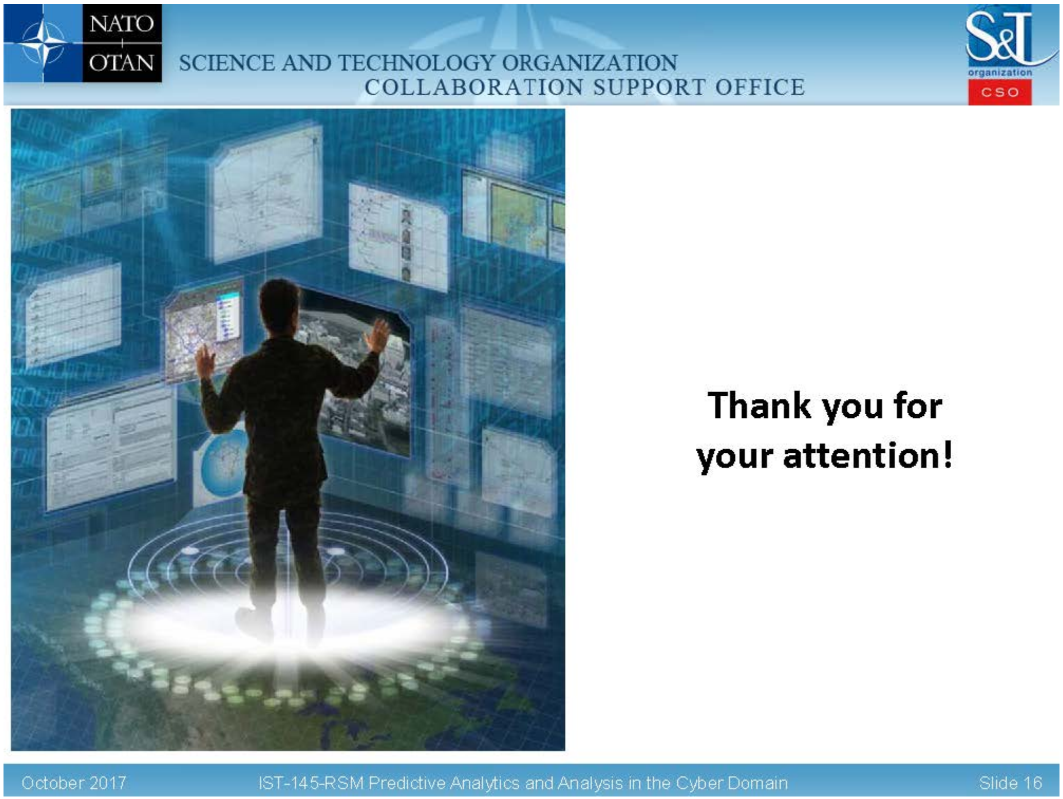
All what you need

S&T events

https://events.sto.nato.int/

PoC
IST Panel Office
Maj. Luc DETIENNE / Mrs. Ayşegül APAYDIN

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 15



NATO OTAN

SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

S&T organization CSO

Thank you for
your attention!

October 2017 IST-145-RSM Predictive Analytics and Analysis in the Cyber Domain Slide 16

NATO OTAN

SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

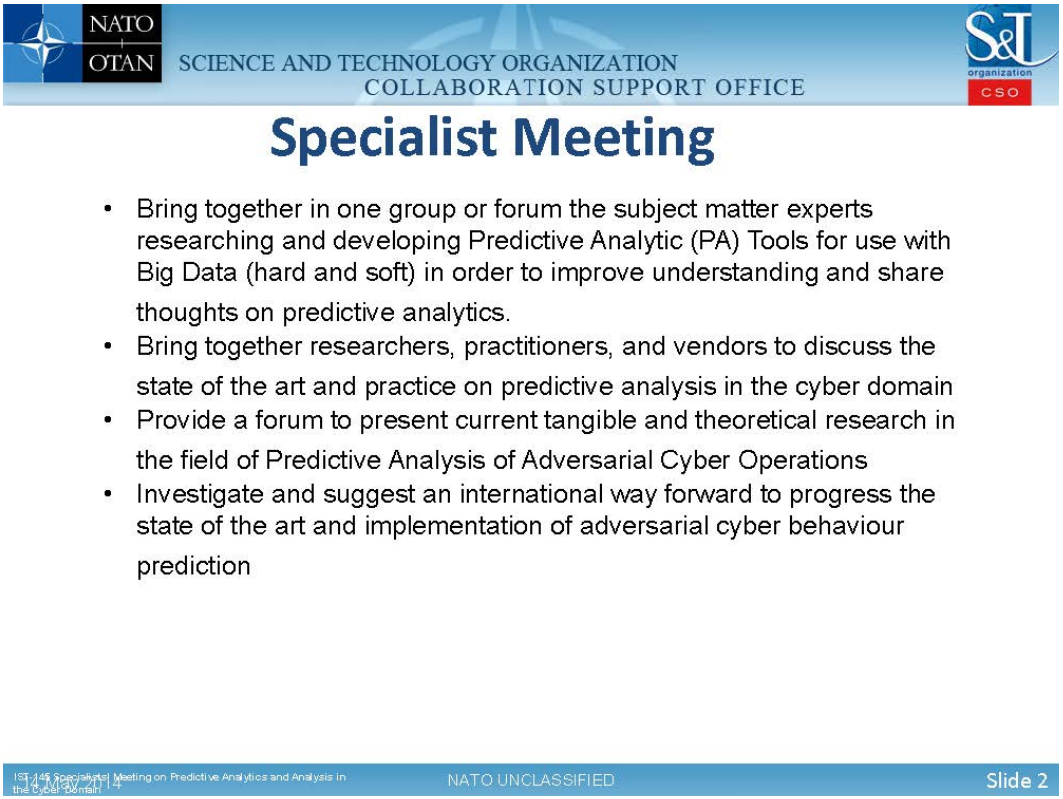
S&T organization CSO

IST-145/RSM-030 Specialists' Meeting

PREDICTIVE ANALYTICS AND ANALYSIS IN THE CYBER DOMAIN

Dr. Dennis McCALLAM
Northrop Grumman
United States
Email: dennis.mccallam@ngc.com

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 1



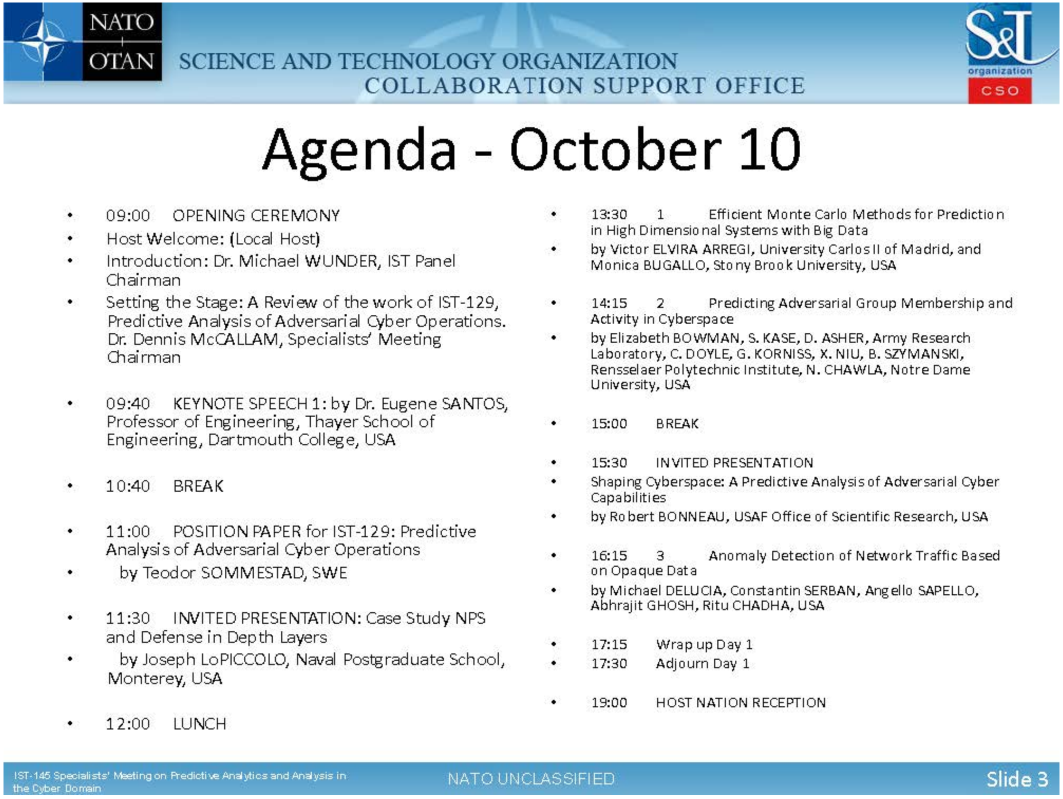
NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

Specialist Meeting

- Bring together in one group or forum the subject matter experts researching and developing Predictive Analytic (PA) Tools for use with Big Data (hard and soft) in order to improve understanding and share thoughts on predictive analytics.
- Bring together researchers, practitioners, and vendors to discuss the state of the art and practice on predictive analysis in the cyber domain
- Provide a forum to present current tangible and theoretical research in the field of Predictive Analysis of Adversarial Cyber Operations
- Investigate and suggest an international way forward to progress the state of the art and implementation of adversarial cyber behaviour prediction

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain 14 May 2014 NATO UNCLASSIFIED Slide 2



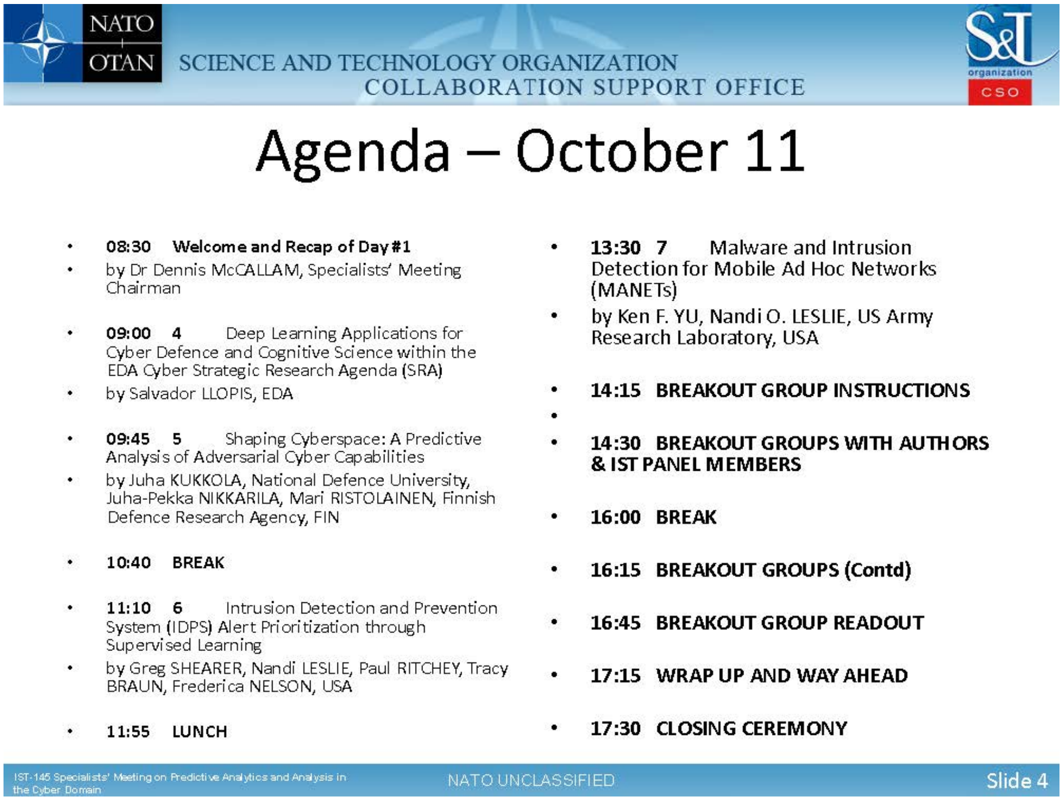
NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

Agenda - October 10

- 09:00 OPENING CEREMONY
- Host Welcome: (Local Host)
- Introduction: Dr. Michael WUNDER, IST Panel Chairman
- Setting the Stage: A Review of the work of IST-129, Predictive Analysis of Adversarial Cyber Operations. Dr. Dennis McCALLAM, Specialists' Meeting Chairman
- 09:40 KEYNOTE SPEECH 1: by Dr. Eugene SANTOS, Professor of Engineering, Thayer School of Engineering, Dartmouth College, USA
- 10:40 BREAK
- 11:00 POSITION PAPER for IST-129: Predictive Analysis of Adversarial Cyber Operations
- by Teodor SOMMESTAD, SWE
- 11:30 INVITED PRESENTATION: Case Study NPS and Defense in Depth Layers
- by Joseph LoPICCOLO, Naval Postgraduate School, Monterey, USA
- 12:00 LUNCH
- 13:30 1 Efficient Monte Carlo Methods for Prediction in High Dimensional Systems with Big Data
- by Victor ELVIRA ARREGI, University Carlos II of Madrid, and Monica BUGALLO, Stony Brook University, USA
- 14:15 2 Predicting Adversarial Group Membership and Activity in Cyberspace
- by Elizabeth BOWMAN, S. KASE, D. ASHER, Army Research Laboratory, C. DOYLE, G. KORNISS, X. NIU, B. SZYMANSKI, Rensselaer Polytechnic Institute, N. CHAWLA, Notre Dame University, USA
- 15:00 BREAK
- 15:30 INVITED PRESENTATION
- Shaping Cyberspace: A Predictive Analysis of Adversarial Cyber Capabilities
- by Robert BONNEAU, USAF Office of Scientific Research, USA
- 16:15 3 Anomaly Detection of Network Traffic Based on Opaque Data
- by Michael DELUCIA, Constantin SERBAN, Angello SAPELLO, Abhrajit GHOSH, Ritu CHADHA, USA
- 17:15 Wrap up Day 1
- 17:30 Adjourn Day 1
- 19:00 HOST NATION RECEPTION

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 3

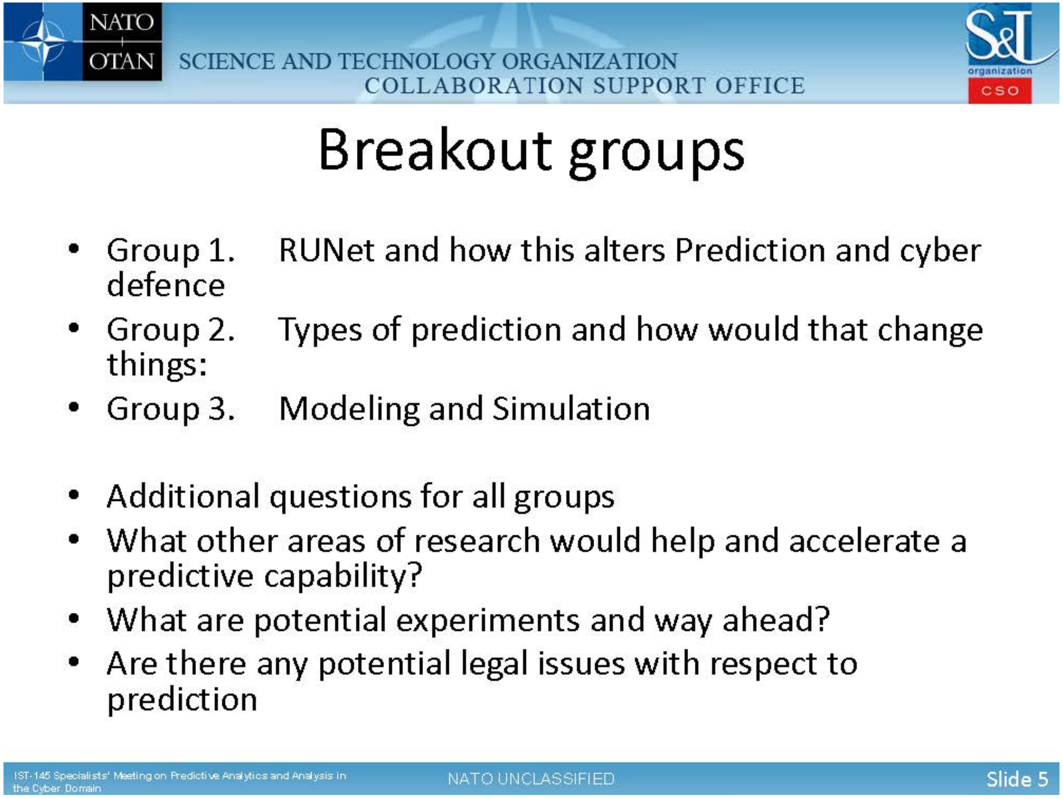

NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

S&T organization CSO

Agenda – October 11

- 08:30 **Welcome and Recap of Day #1**
- by Dr Dennis McCALLAM, Specialists' Meeting Chairman
- 09:00 4 Deep Learning Applications for Cyber Defence and Cognitive Science within the EDA Cyber Strategic Research Agenda (SRA)
- by Salvador LLOPIS, EDA
- 09:45 5 Shaping Cyberspace: A Predictive Analysis of Adversarial Cyber Capabilities
- by Juha KUKKOLA, National Defence University, Juha-Pekka NIKKARILA, Mari RISTOLAINEN, Finnish Defence Research Agency, FIN
- 10:40 BREAK
- 11:10 6 Intrusion Detection and Prevention System (IDPS) Alert Prioritization through Supervised Learning
- by Greg SHEARER, Nandi LESLIE, Paul RITCHEY, Tracy BRAUN, Frederica NELSON, USA
- 11:55 LUNCH
- 13:30 7 Malware and Intrusion Detection for Mobile Ad Hoc Networks (MANETs)
- by Ken F. YU, Nandi O. LESLIE, US Army Research Laboratory, USA
- **14:15 BREAKOUT GROUP INSTRUCTIONS**
- **14:30 BREAKOUT GROUPS WITH AUTHORS & IST PANEL MEMBERS**
- **16:00 BREAK**
- **16:15 BREAKOUT GROUPS (Contd)**
- **16:45 BREAKOUT GROUP READOUT**
- **17:15 WRAP UP AND WAY AHEAD**
- **17:30 CLOSING CEREMONY**

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 4


NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

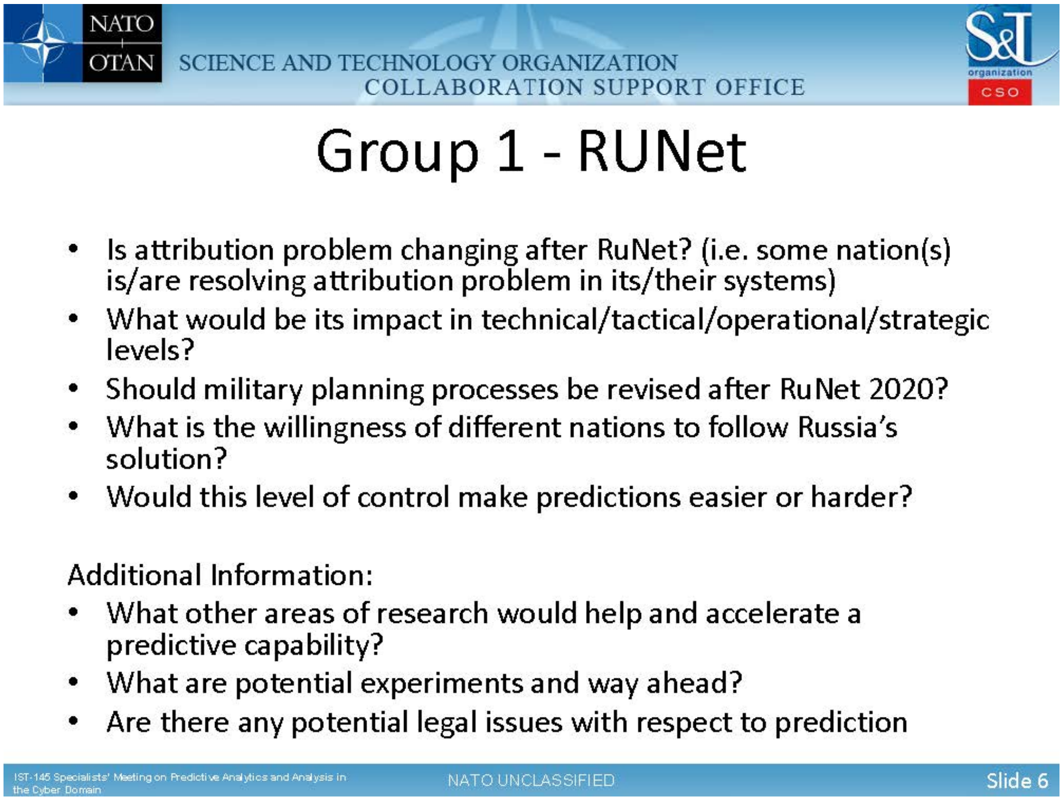
S&T organization CSO

Breakout groups

- Group 1. RUNet and how this alters Prediction and cyber defence
- Group 2. Types of prediction and how would that change things:
- Group 3. Modeling and Simulation

- Additional questions for all groups
- What other areas of research would help and accelerate a predictive capability?
- What are potential experiments and way ahead?
- Are there any potential legal issues with respect to prediction

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 5



NATO OTAN

SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

S&T organization CSO

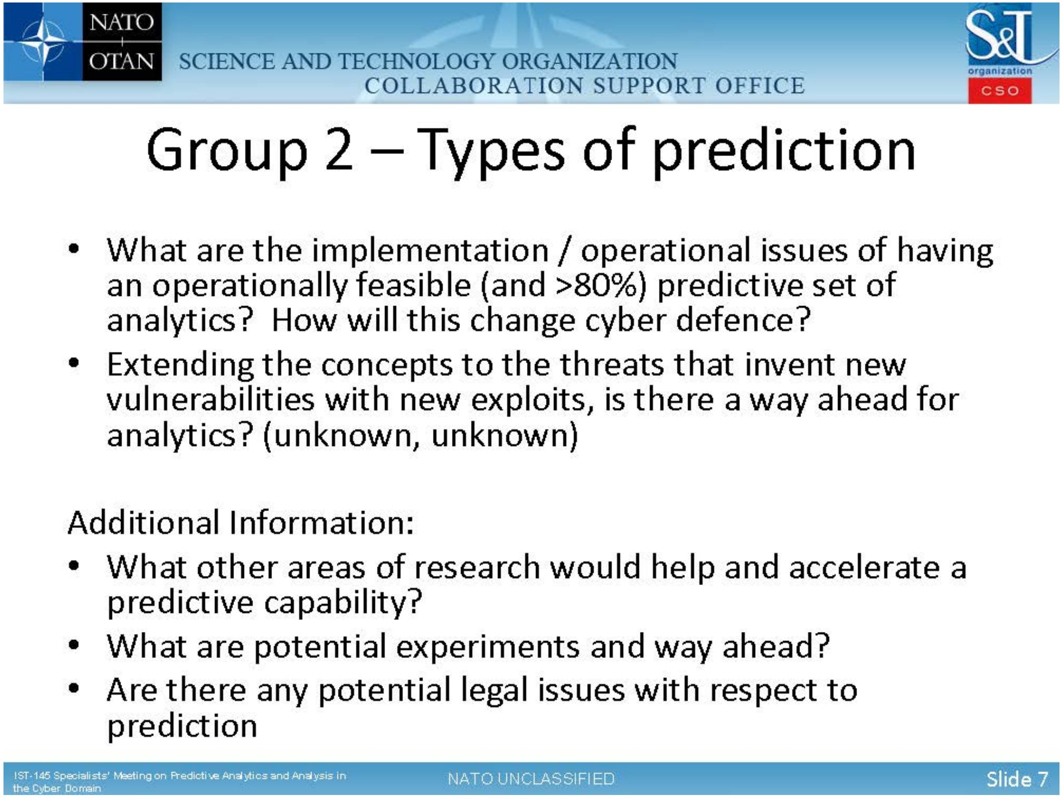
Group 1 - RUNet

- Is attribution problem changing after RuNet? (i.e. some nation(s) is/are resolving attribution problem in its/their systems)
- What would be its impact in technical/tactical/operational/strategic levels?
- Should military planning processes be revised after RuNet 2020?
- What is the willingness of different nations to follow Russia's solution?
- Would this level of control make predictions easier or harder?

Additional Information:

- What other areas of research would help and accelerate a predictive capability?
- What are potential experiments and way ahead?
- Are there any potential legal issues with respect to prediction

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 6



NATO OTAN

SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

S&T organization CSO

Group 2 – Types of prediction

- What are the implementation / operational issues of having an operationally feasible (and >80%) predictive set of analytics? How will this change cyber defence?
- Extending the concepts to the threats that invent new vulnerabilities with new exploits, is there a way ahead for analytics? (unknown, unknown)

Additional Information:

- What other areas of research would help and accelerate a predictive capability?
- What are potential experiments and way ahead?
- Are there any potential legal issues with respect to prediction

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 7

Group 3 – Modeling & Simulation

- What are the Modelling and Simulation requirements to adequately testing/developing predictive systems?
- If predictive analytics were used to identify attacks, could that be constantly run to not only identify potential attacks, but to generate the correct patches (rendering things the known-known signature based situation). How would that be done? How does that change cyber defence as we know it?
- Should this be done in real time?

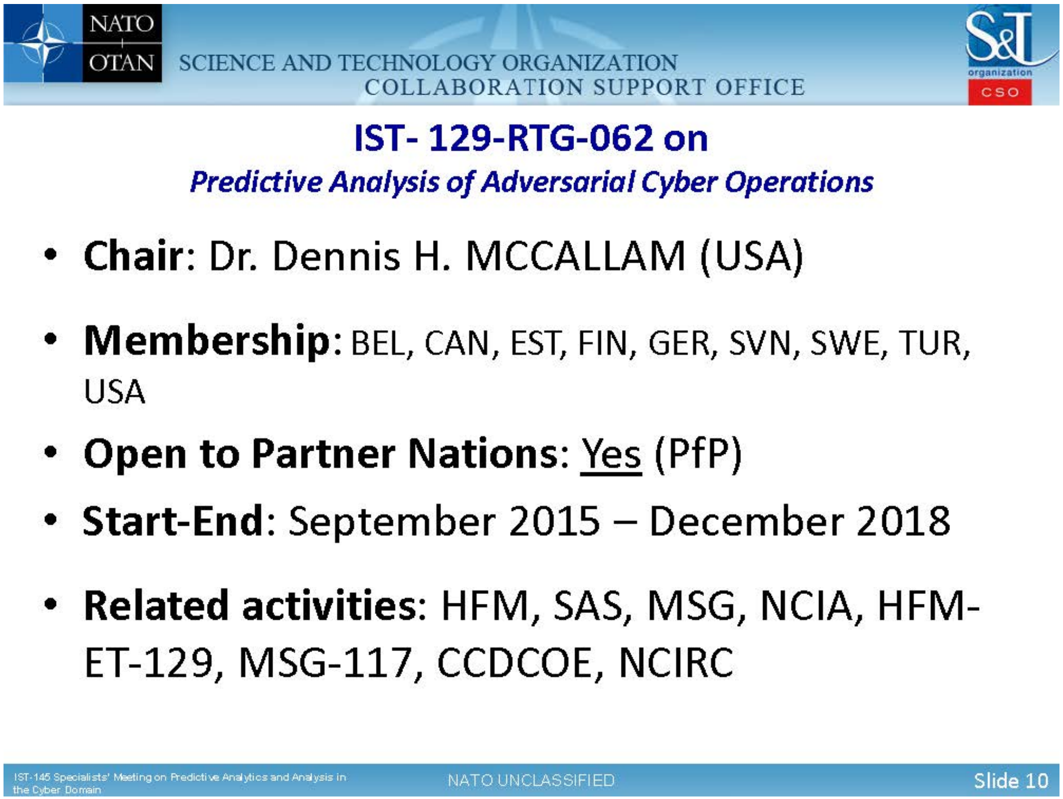
Additional Information:

- What other areas of research would help and accelerate a predictive capability?
- What are potential experiments and way ahead?
- Are there any potential legal issues with respect to prediction

IST-129
Predictive Analysis of Adversarial Cyber Operations

Predictive analysis has to consider the past and present to properly predict the future

Chair: Dr. Dennis McCallam USA



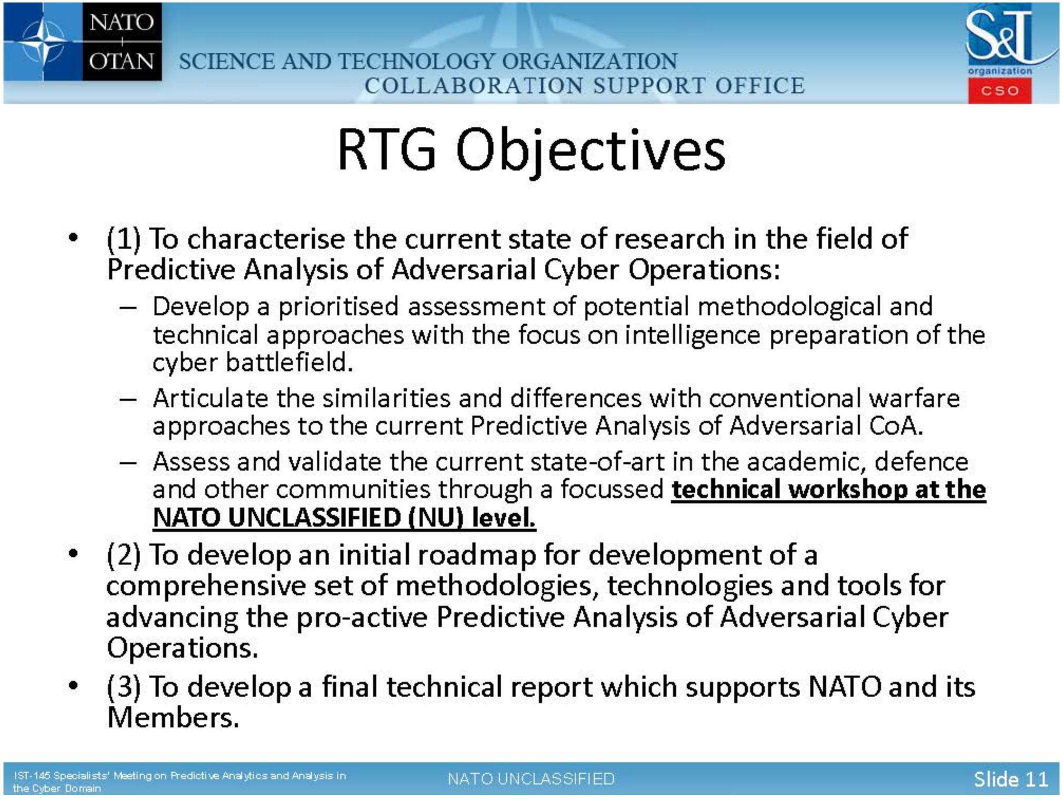
NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

IST- 129-RTG-062 on

Predictive Analysis of Adversarial Cyber Operations

- **Chair**: Dr. Dennis H. MCCALLAM (USA)
- **Membership**: BEL, CAN, EST, FIN, GER, SVN, SWE, TUR, USA
- **Open to Partner Nations**: Yes (PfP)
- **Start-End**: September 2015 – December 2018
- **Related activities**: HFM, SAS, MSG, NCIA, HFM-ET-129, MSG-117, CCDCOE, NCIRC

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 10

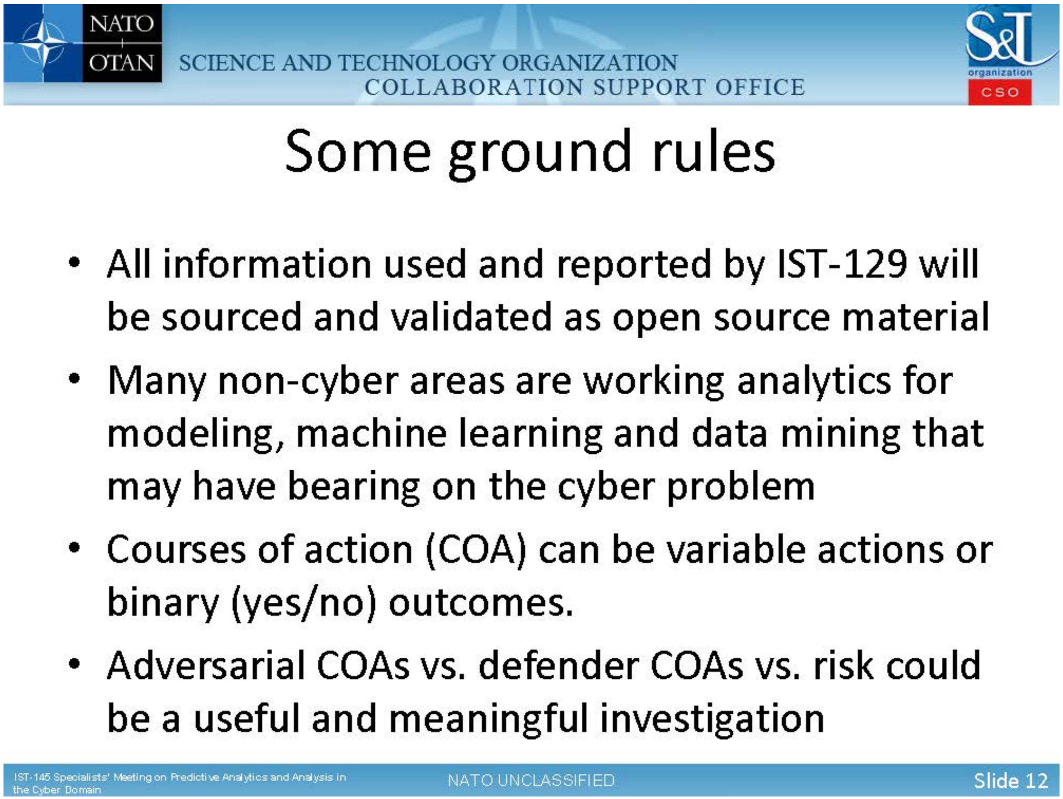


NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

RTG Objectives

- (1) To characterise the current state of research in the field of Predictive Analysis of Adversarial Cyber Operations:
  - Develop a prioritised assessment of potential methodological and technical approaches with the focus on intelligence preparation of the cyber battlefield.
  - Articulate the similarities and differences with conventional warfare approaches to the current Predictive Analysis of Adversarial CoA.
  - Assess and validate the current state-of-art in the academic, defence and other communities through a focussed **technical workshop at the NATO UNCLASSIFIED (NU) level.**
- (2) To develop an initial roadmap for development of a comprehensive set of methodologies, technologies and tools for advancing the pro-active Predictive Analysis of Adversarial Cyber Operations.
- (3) To develop a final technical report which supports NATO and its Members.

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 11



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

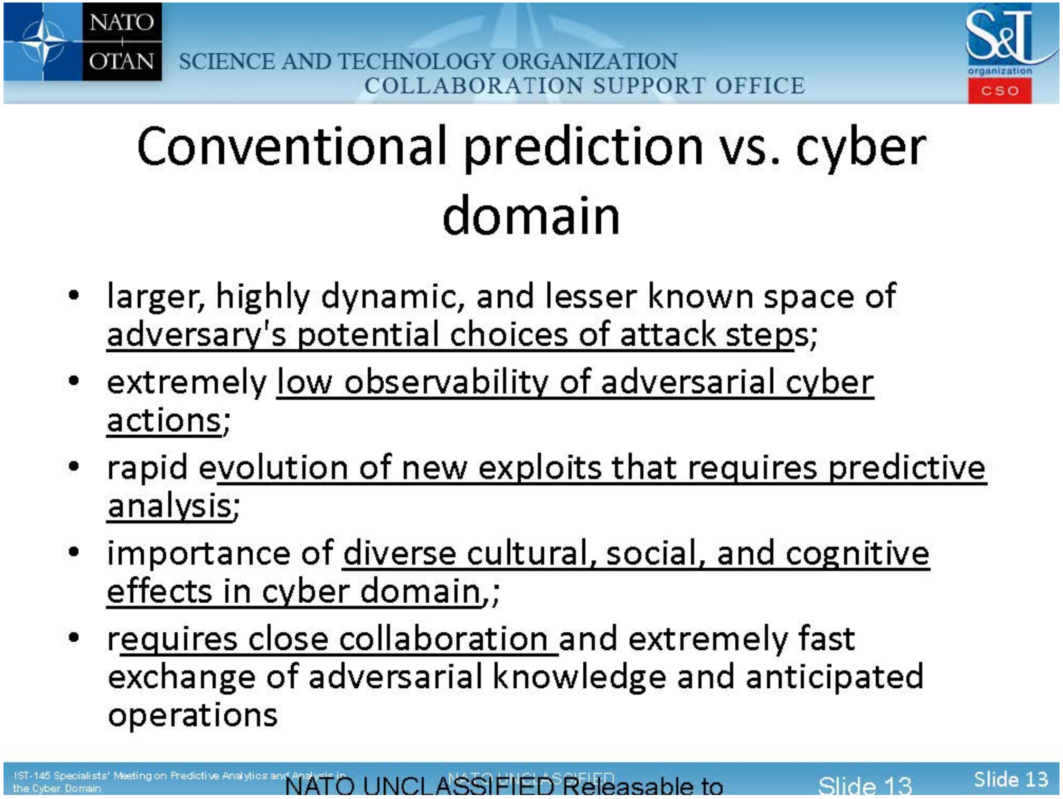
Some ground rules

- All information used and reported by IST-129 will be sourced and validated as open source material
- Many non-cyber areas are working analytics for modeling, machine learning and data mining that may have bearing on the cyber problem
- Courses of action (COA) can be variable actions or binary (yes/no) outcomes.
- Adversarial COAs vs. defender COAs vs. risk could be a useful and meaningful investigation

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain

NATO UNCLASSIFIED

Slide 12



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

S&T organization CSO

Conventional prediction vs. cyber domain

- larger, highly dynamic, and lesser known space of adversary's potential choices of attack steps;
- extremely low observability of adversarial cyber actions;
- rapid evolution of new exploits that requires predictive analysis;
- importance of diverse cultural, social, and cognitive effects in cyber domain,;
- requires close collaboration and extremely fast exchange of adversarial knowledge and anticipated operations

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain

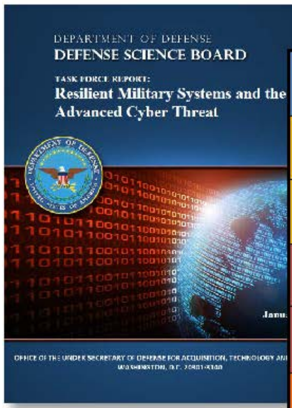
NATO UNCLASSIFIED

NATO UNCLASSIFIED Releasable to

Slide 13

Slide 13

A useful context for the Threat



DEPARTMENT OF DEFENSE
DEFENSE SCIENCE BOARD
TASK FORCE REPORT:
Resilient Military Systems and the Advanced Cyber Threat

Janu

OFFICE OF THE UNDER SECRETARY OF DEFENSE FOR ACQUISITION, TECHNOLOGY AN
WASHINGTON, D.C. 20301-3140

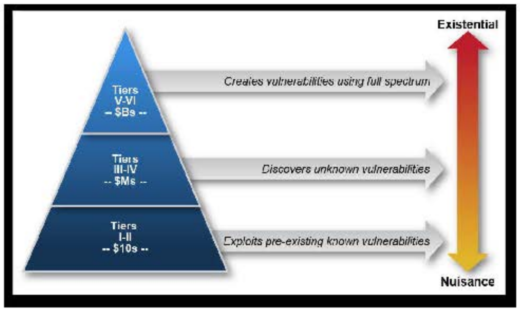
Threat Tier	Description
I	*Practitioners who rely on others to develop the malicious code.*
II	*Practitioners with the ability to develop their own tools (from publically known vulnerabilities).*
III	*Practitioners, who focus on the discovery and use of unknown malicious code.*
IV	*Criminal or state actors who are organized, highly technical, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.*
V	*State actors who create supply chain vulnerabilities.*
VI	*States with the ability to successfully execute full spectrum cyber operations*

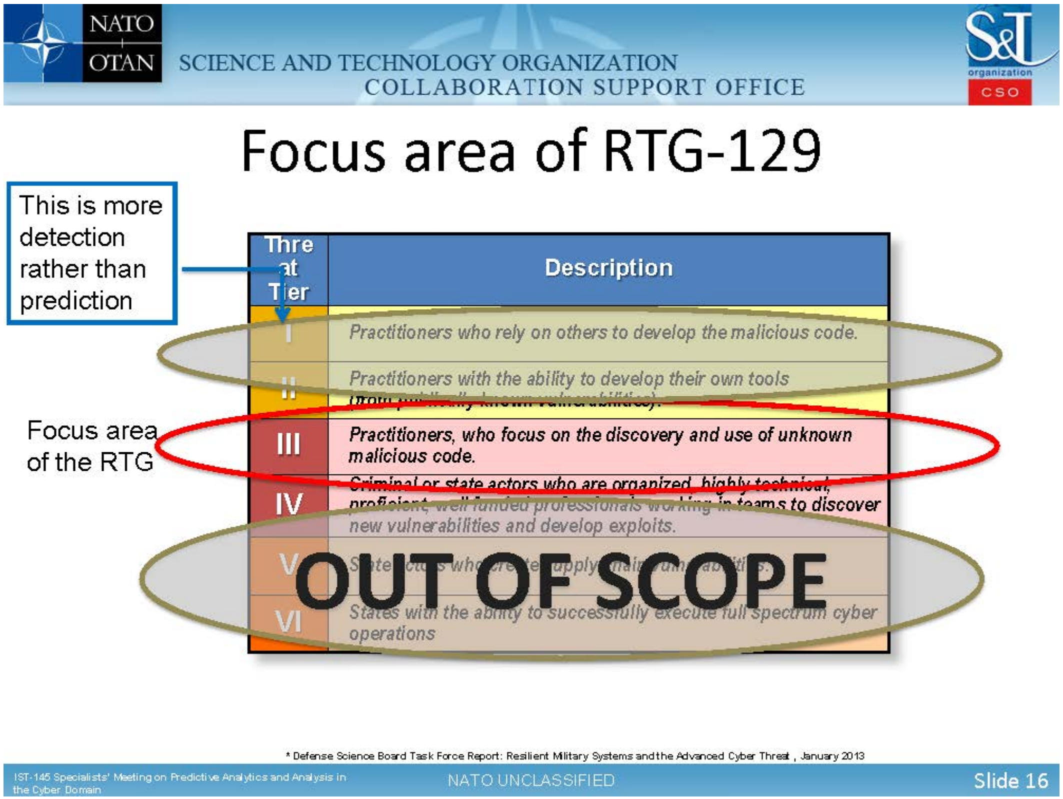
How does this context for threats relate to or steer predictive analysis? Is there a relationship between Threat Tier and analytic type?

* Defense Science Board Task Force Report: Resilient Military Systems and the Advanced Cyber Threat , January 2013

Capability equates to investment

- Tiers I and II attackers primarily *exploit known* vulnerabilities
- Tiers III and IV attackers are better funded and have a level of expertise and sophistication sufficient to *discover new* vulnerabilities in systems and to exploit them
- Tiers V and VI attackers can invest large amounts of money (billions) and time (years) to *actually create* vulnerabilities in systems, including systems that are otherwise strongly protected.


Existential
Tiers
V-VI
-- $Bs --
Creates vulnerabilities using full spectrum
Tiers
III-IV
-- $Ms --
Discovers unknown vulnerabilities
Tiers
I-II
-- $10s --
Exploits pre-existing known vulnerabilities
Nuisance



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

Focus area of RTG-129

This is more detection rather than prediction

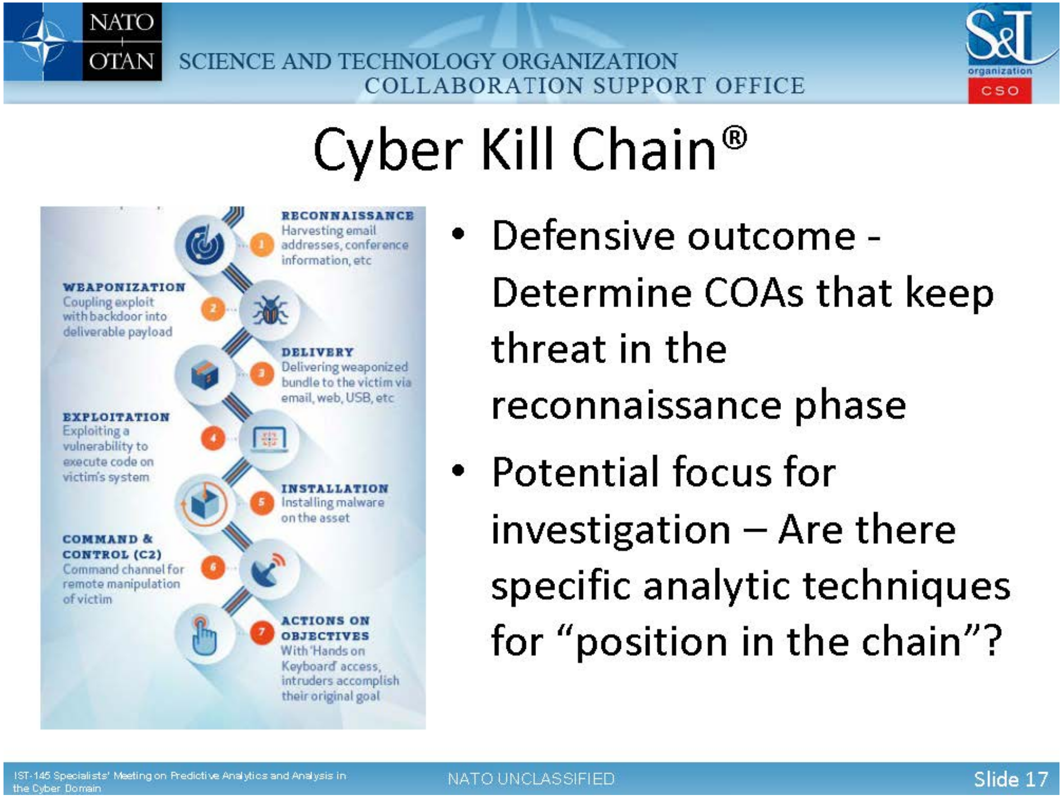
Focus area of the RTG

Threat Tier	Description
I	Practitioners who rely on others to develop the malicious code.
II	Practitioners with the ability to develop their own tools (from publically known vulnerabilities).
III	Practitioners, who focus on the discovery and use of unknown malicious code.
IV	Criminal or state actors who are organized, highly technical, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.
V	State actors who create/apply [illegible] abilities.
VI	States with the ability to successfully execute full spectrum cyber operations

OUT OF SCOPE

* Defense Science Board Task Force Report: Resilient Military Systems and the Advanced Cyber Threat , January 2013

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 16



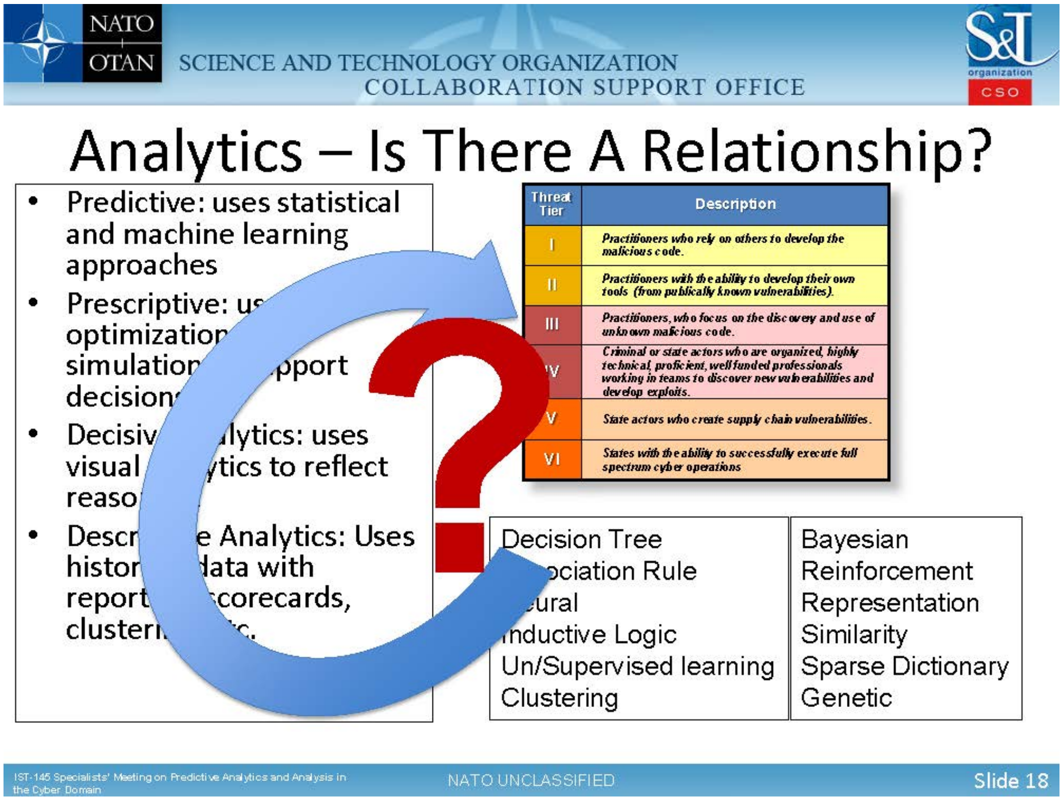
NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

Cyber Kill Chain®

RECONNAISSANCE
Harvesting email addresses, conference information, etc
1
WEAPONIZATION
Coupling exploit with backdoor into deliverable payload
2
DELIVERY
Delivering weaponized bundle to the victim via email, web, USB, etc
3
EXPLOITATION
Exploiting a vulnerability to execute code on victim's system
4
INSTALLATION
Installing malware on the asset
5
COMMAND & CONTROL (C2)
Command channel for remote manipulation of victim
6
ACTIONS ON OBJECTIVES
With 'Hands on Keyboard' access, intruders accomplish their original goal
7

- Defensive outcome - Determine COAs that keep threat in the reconnaissance phase
- Potential focus for investigation – Are there specific analytic techniques for "position in the chain"?

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 17



NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

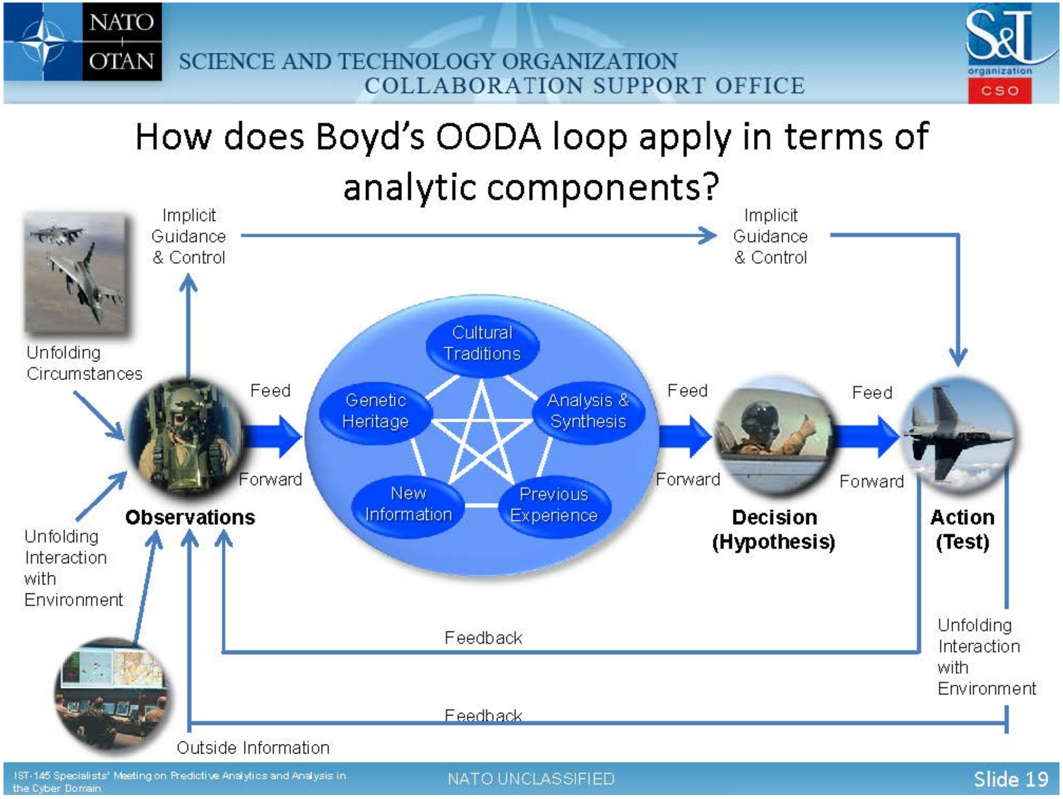
Analytics – Is There A Relationship?

- Predictive: uses statistical and machine learning approaches
- Prescriptive: us optimization simulation pport decision
- Decisiv lytics: uses visual ytics to reflect reaso
- Descr e Analytics: Uses histor data with report corecards, clusteri c.

Threat Tier	Description
I	*Practitioners who rely on others to develop the malicious code.*
II	*Practitioners with the ability to develop their own tools (from publically known vulnerabilities).*
III	*Practitioners, who focus on the discovery and use of unknown malicious code.*
IV	*Criminal or state actors who are organized, highly technical, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.*
V	*State actors who create supply chain vulnerabilities.*
VI	*States with the ability to successfully execute full spectrum cyber operations*

Decision Tree
ociation Rule
eural
nductive Logic
Un/Supervised learning
Clustering

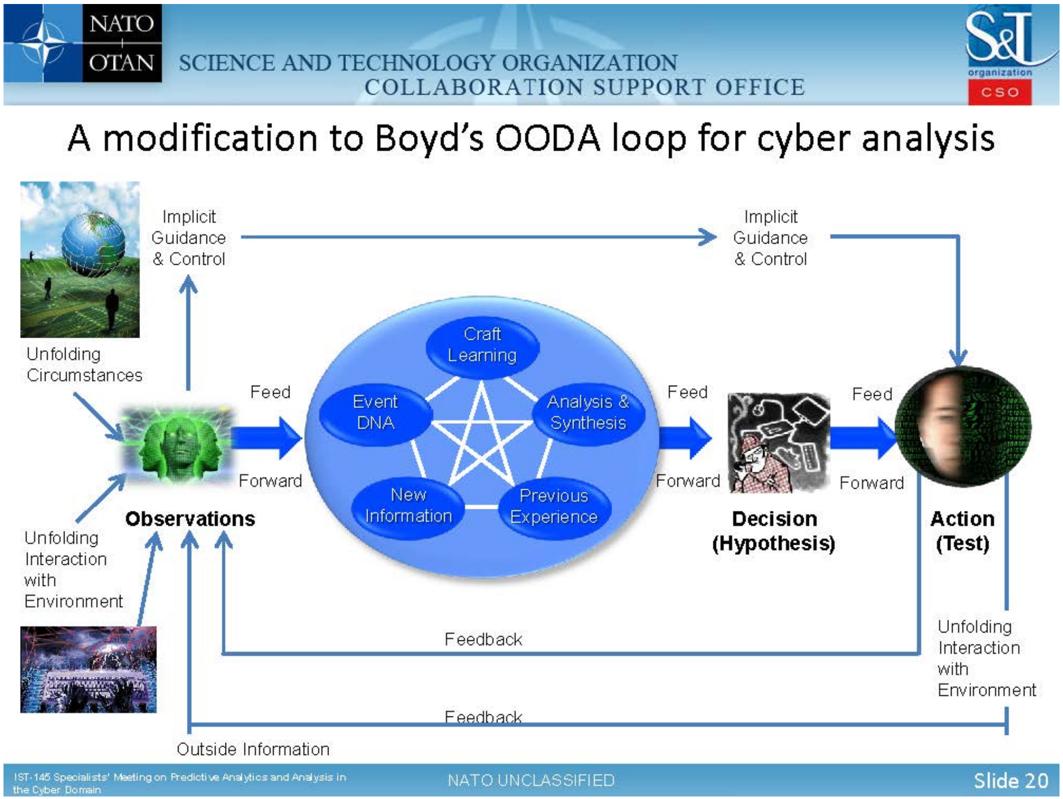
Bayesian
Reinforcement
Representation
Similarity
Sparse Dictionary
Genetic



NATO
OTAN
SCIENCE AND TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

How does Boyd's OODA loop apply in terms of analytic components?

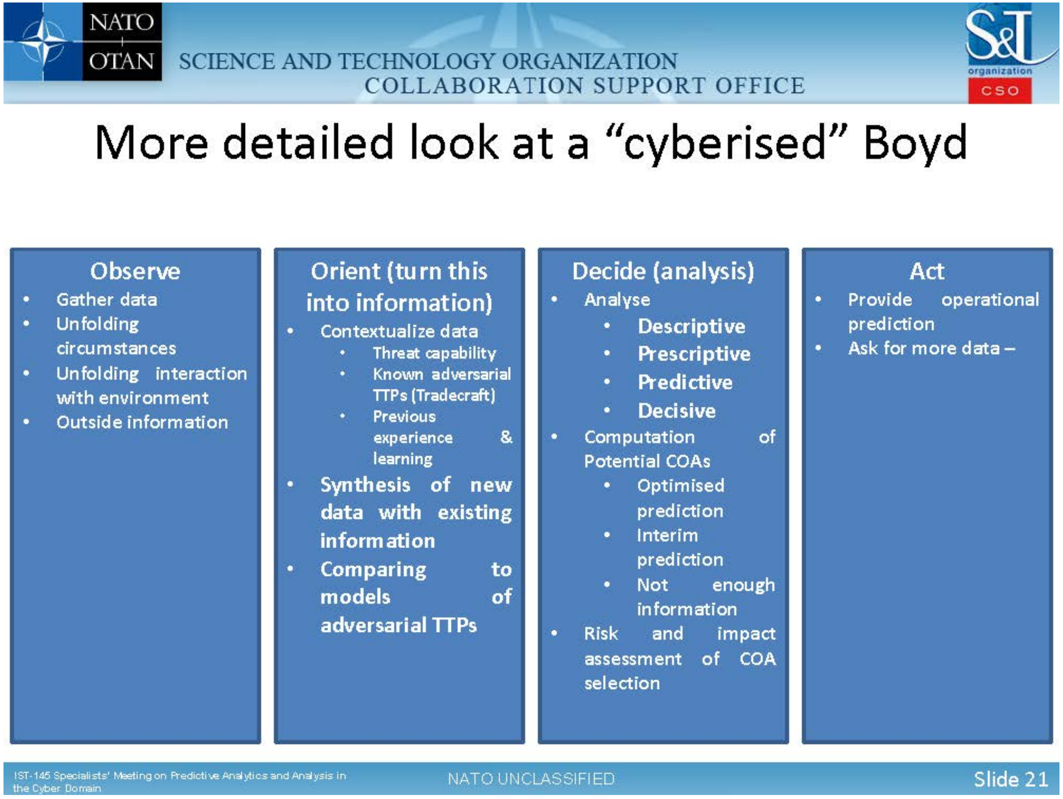
Implicit Guidance & Control
Implicit Guidance & Control
Unfolding Circumstances
Feed
Forward
Cultural Traditions
Genetic Heritage
Analysis & Synthesis
New Information
Previous Experience
Feed
Forward
Feed
Forward
Unfolding Interaction with Environment
Observations
Decision (Hypothesis)
Action (Test)
Unfolding Interaction with Environment
Feedback
Feedback
Outside Information



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

A modification to Boyd's OODA loop for cyber analysis

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 20



NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

More detailed look at a "cyberised" Boyd

Observe

- Gather data
- Unfolding circumstances
- Unfolding interaction with environment
- Outside information

Orient (turn this into information)

- Contextualize data
  - Threat capability
  - Known adversarial TTPs (Tradecraft)
  - Previous experience & learning
- **Synthesis of new data with existing information**
- **Comparing to models of adversarial TTPs**

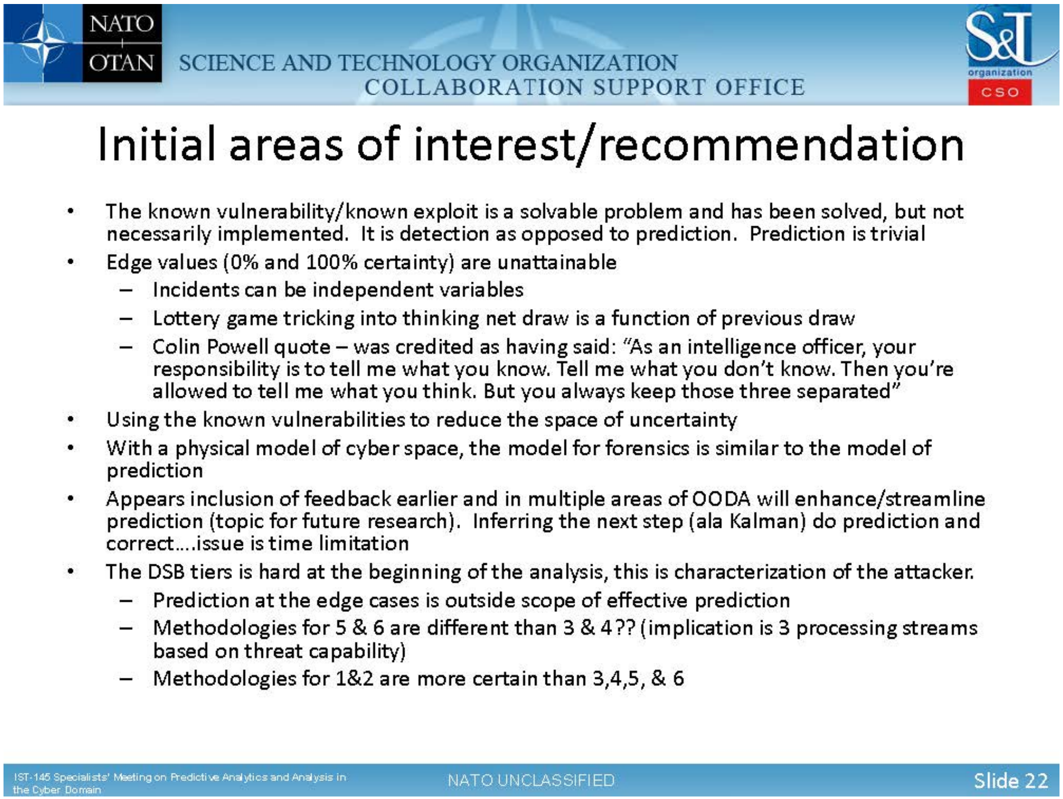
Decide (analysis)

- Analyse
  - **Descriptive**
  - **Prescriptive**
  - **Predictive**
  - **Decisive**
- Computation of Potential COAs
  - Optimised prediction
  - Interim prediction
  - Not enough information
- Risk and impact assessment of COA selection

Act

- Provide operational prediction
- Ask for more data –

IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 21

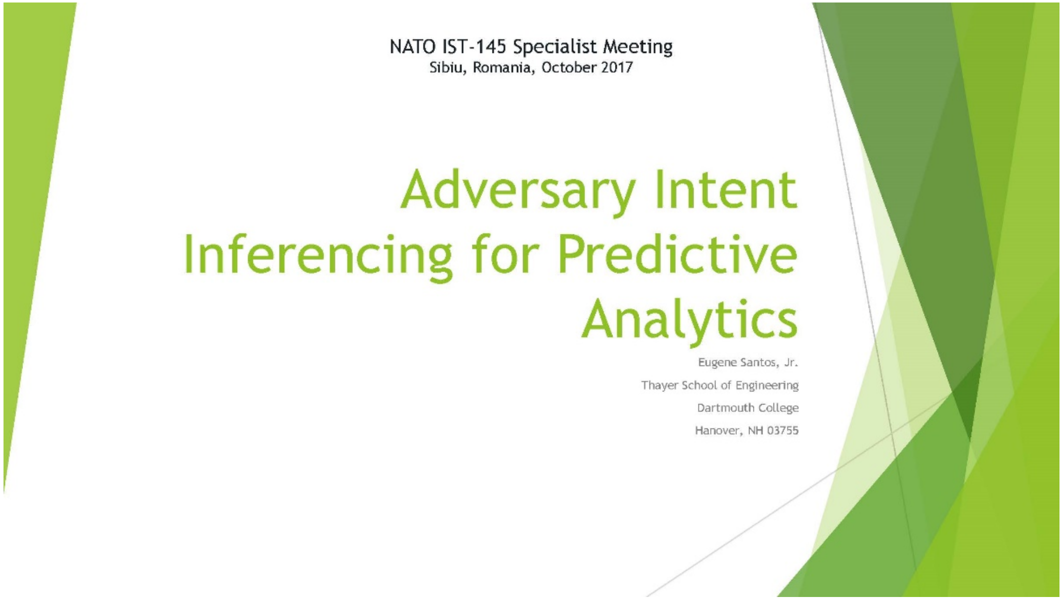


NATO OTAN SCIENCE AND TECHNOLOGY ORGANIZATION COLLABORATION SUPPORT OFFICE

Initial areas of interest/recommendation

- The known vulnerability/known exploit is a solvable problem and has been solved, but not necessarily implemented. It is detection as opposed to prediction. Prediction is trivial
- Edge values (0% and 100% certainty) are unattainable
  - Incidents can be independent variables
  - Lottery game tricking into thinking net draw is a function of previous draw
  - Colin Powell quote – was credited as having said: "As an intelligence officer, your responsibility is to tell me what you know. Tell me what you don't know. Then you're allowed to tell me what you think. But you always keep those three separated"
- Using the known vulnerabilities to reduce the space of uncertainty
- With a physical model of cyber space, the model for forensics is similar to the model of prediction
- Appears inclusion of feedback earlier and in multiple areas of OODA will enhance/streamline prediction (topic for future research). Inferring the next step (ala Kalman) do prediction and correct….issue is time limitation
- The DSB tiers is hard at the beginning of the analysis, this is characterization of the attacker.
  - Prediction at the edge cases is outside scope of effective prediction
  - Methodologies for 5 & 6 are different than 3 & 4?? (implication is 3 processing streams based on threat capability)
  - Methodologies for 1&2 are more certain than 3,4,5, & 6

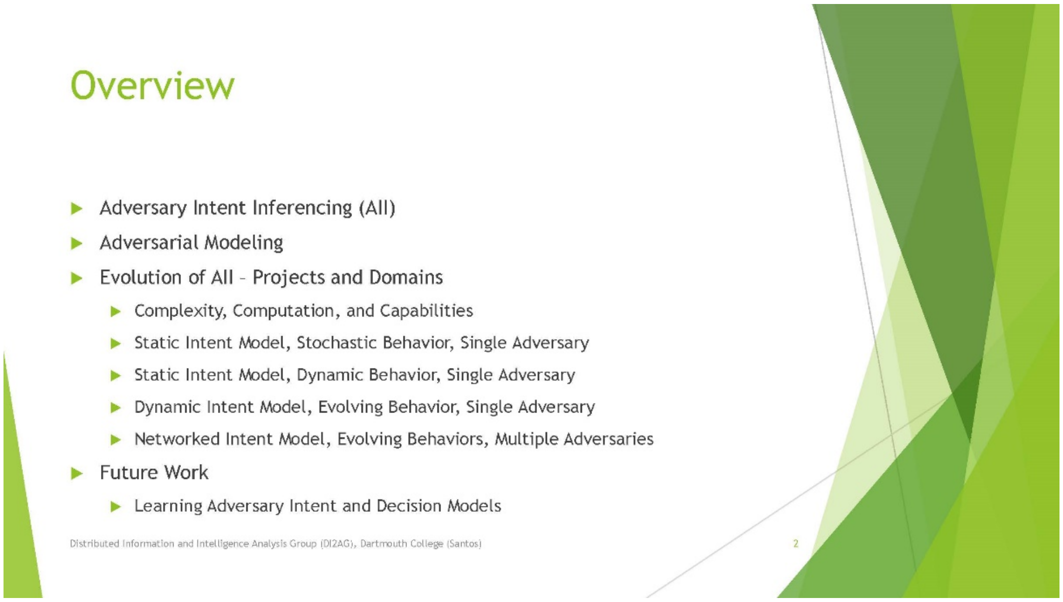
IST-145 Specialists' Meeting on Predictive Analytics and Analysis in the Cyber Domain NATO UNCLASSIFIED Slide 22



NATO IST-145 Specialist Meeting
Sibiu, Romania, October 2017

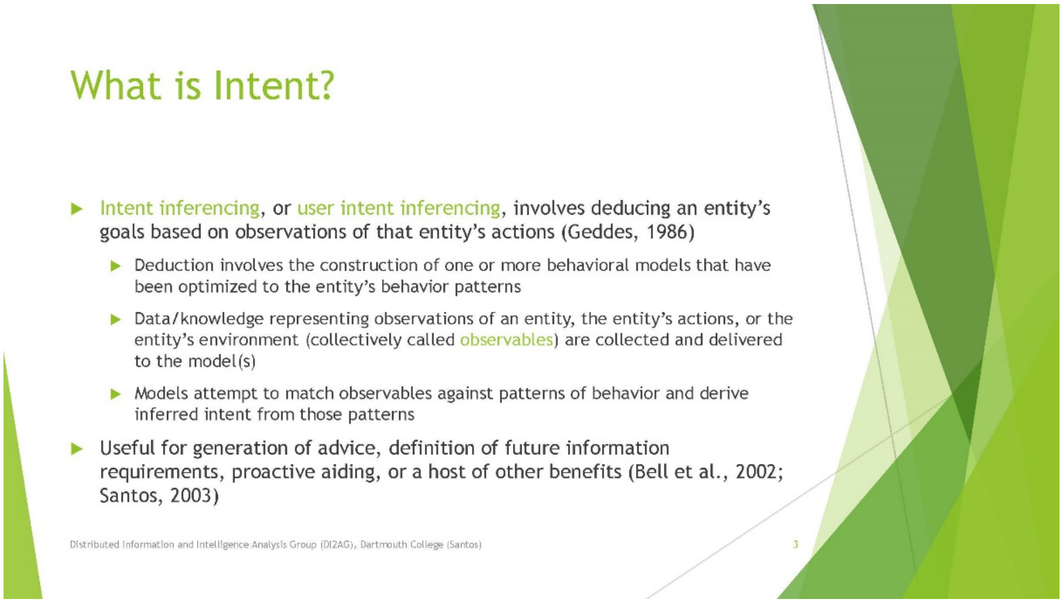
Adversary Intent Inferencing for Predictive Analytics

Eugene Santos, Jr.
Thayer School of Engineering
Dartmouth College
Hanover, NH 03755



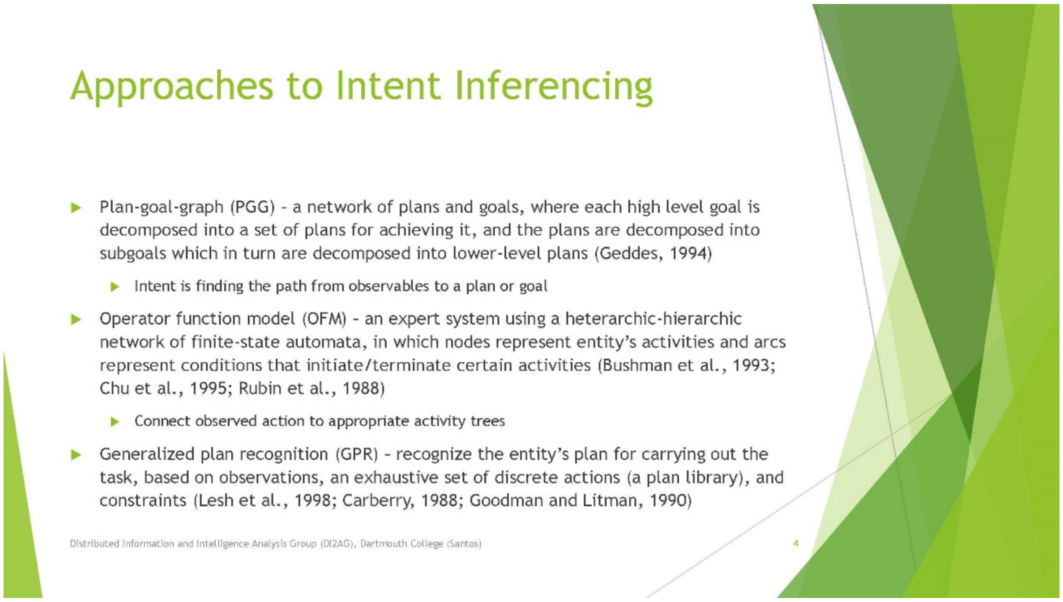
Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models



What is Intent?

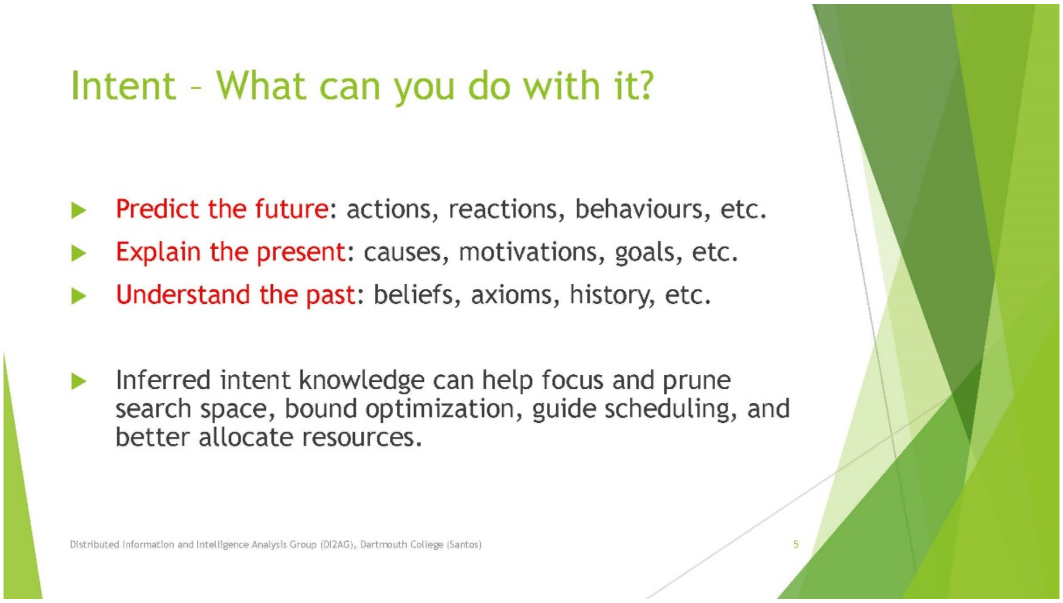
- Intent inferencing, or user intent inferencing, involves deducing an entity's goals based on observations of that entity's actions (Geddes, 1986)
  - Deduction involves the construction of one or more behavioral models that have been optimized to the entity's behavior patterns
  - Data/knowledge representing observations of an entity, the entity's actions, or the entity's environment (collectively called observables) are collected and delivered to the model(s)
  - Models attempt to match observables against patterns of behavior and derive inferred intent from those patterns
- Useful for generation of advice, definition of future information requirements, proactive aiding, or a host of other benefits (Bell et al., 2002; Santos, 2003)



Approaches to Intent Inferencing

- Plan-goal-graph (PGG) – a network of plans and goals, where each high level goal is decomposed into a set of plans for achieving it, and the plans are decomposed into subgoals which in turn are decomposed into lower-level plans (Geddes, 1994)
  - Intent is finding the path from observables to a plan or goal
- Operator function model (OFM) – an expert system using a heterarchic-hierarchic network of finite-state automata, in which nodes represent entity's activities and arcs represent conditions that initiate/terminate certain activities (Bushman et al., 1993; Chu et al., 1995; Rubin et al., 1988)
  - Connect observed action to appropriate activity trees
- Generalized plan recognition (GPR) – recognize the entity's plan for carrying out the task, based on observations, an exhaustive set of discrete actions (a plan library), and constraints (Lesh et al., 1998; Carberry, 1988; Goodman and Litman, 1990)

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 4

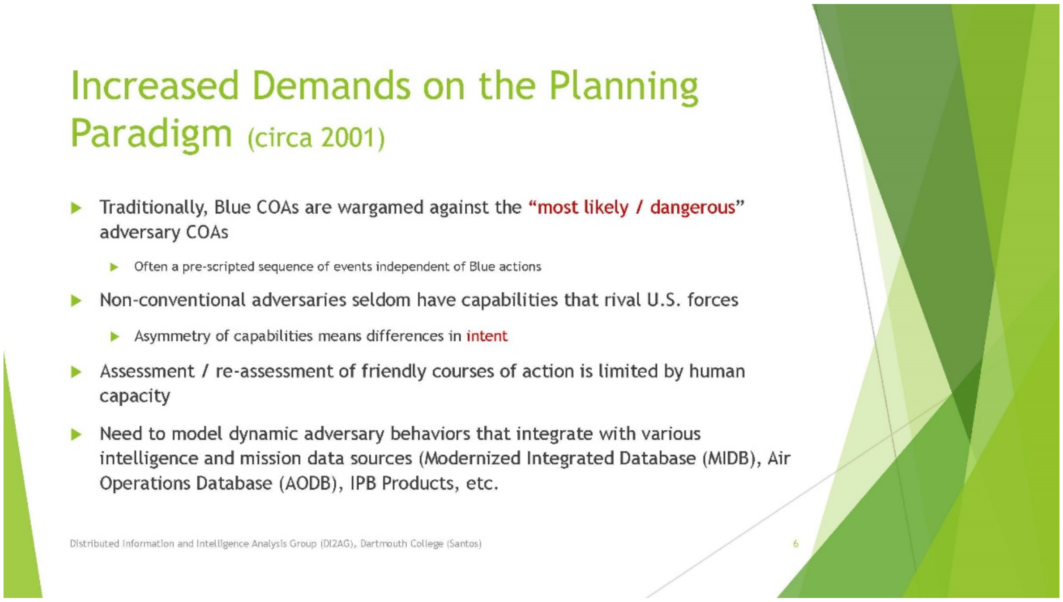


Intent – What can you do with it?

- Predict the future: actions, reactions, behaviours, etc.
- Explain the present: causes, motivations, goals, etc.
- Understand the past: beliefs, axioms, history, etc.

- Inferred intent knowledge can help focus and prune search space, bound optimization, guide scheduling, and better allocate resources.

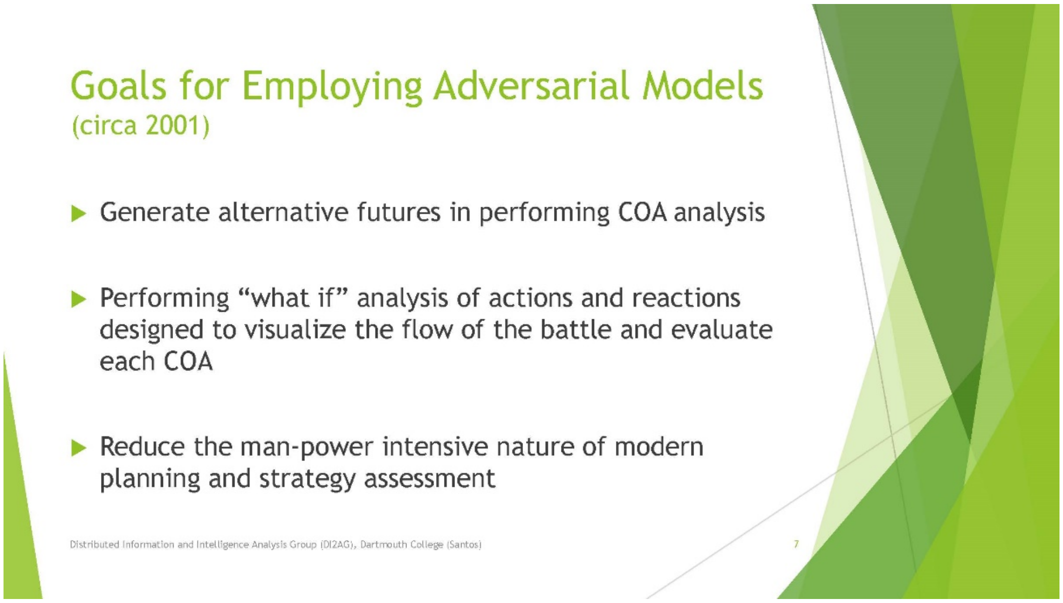
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 5



Increased Demands on the Planning Paradigm (circa 2001)

- Traditionally, Blue COAs are wargamed against the "most likely / dangerous" adversary COAs
  - Often a pre-scripted sequence of events independent of Blue actions
- Non-conventional adversaries seldom have capabilities that rival U.S. forces
  - Asymmetry of capabilities means differences in intent
- Assessment / re-assessment of friendly courses of action is limited by human capacity
- Need to model dynamic adversary behaviors that integrate with various intelligence and mission data sources (Modernized Integrated Database (MIDB), Air Operations Database (AODB), IPB Products, etc.

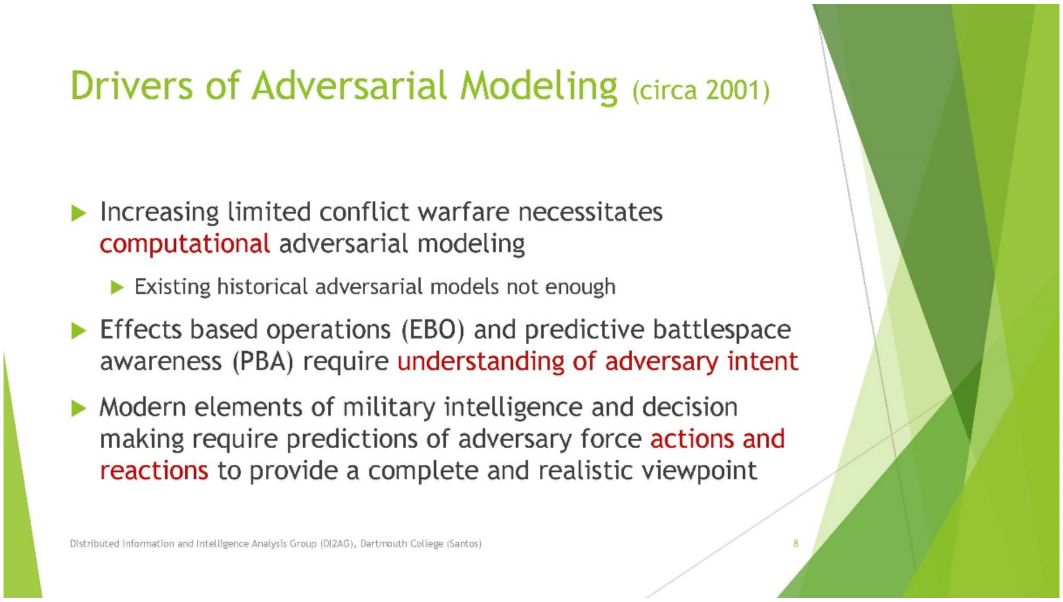
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 6



Goals for Employing Adversarial Models (circa 2001)

- Generate alternative futures in performing COA analysis
- Performing "what if" analysis of actions and reactions designed to visualize the flow of the battle and evaluate each COA
- Reduce the man-power intensive nature of modern planning and strategy assessment

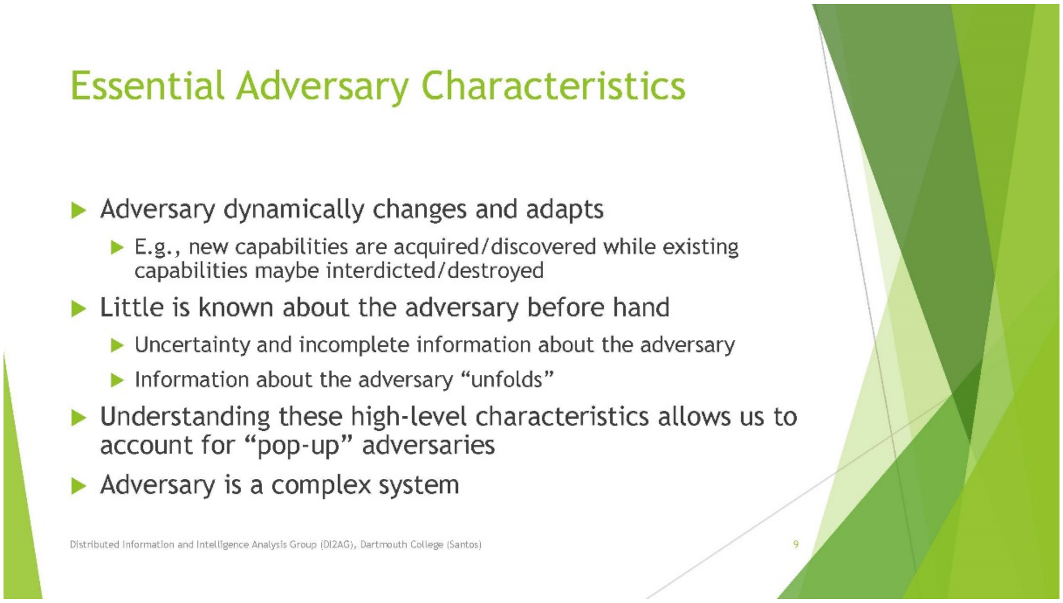
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 7



Drivers of Adversarial Modeling (circa 2001)

- Increasing limited conflict warfare necessitates computational adversarial modeling
  - Existing historical adversarial models not enough
- Effects based operations (EBO) and predictive battlespace awareness (PBA) require understanding of adversary intent
- Modern elements of military intelligence and decision making require predictions of adversary force actions and reactions to provide a complete and realistic viewpoint

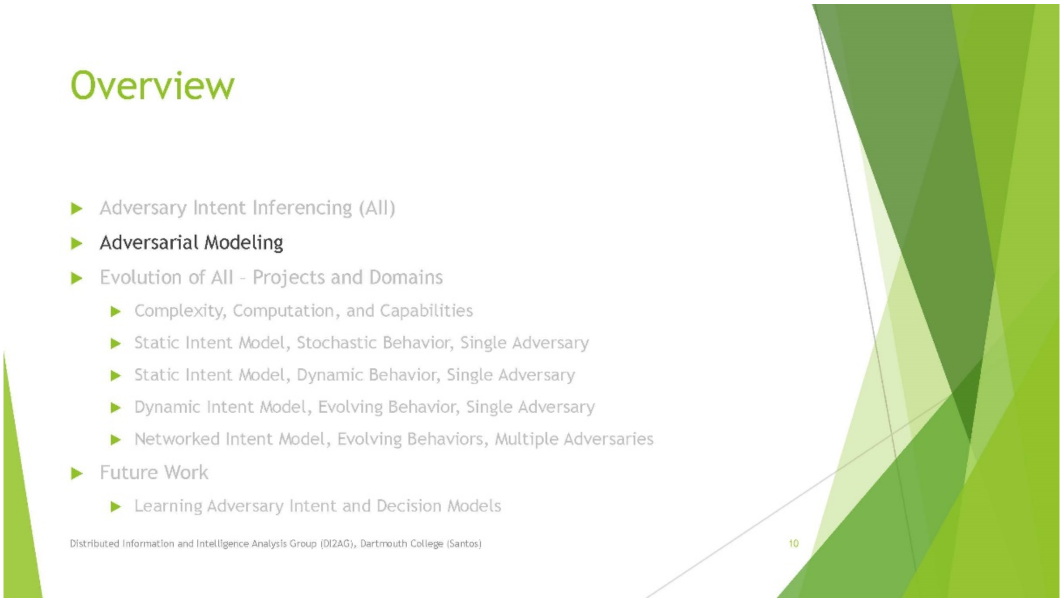
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 8



Essential Adversary Characteristics

- Adversary dynamically changes and adapts
  - E.g., new capabilities are acquired/discovered while existing capabilities maybe interdicted/destroyed
- Little is known about the adversary before hand
  - Uncertainty and incomplete information about the adversary
  - Information about the adversary "unfolds"
- Understanding these high-level characteristics allows us to account for "pop-up" adversaries
- Adversary is a complex system

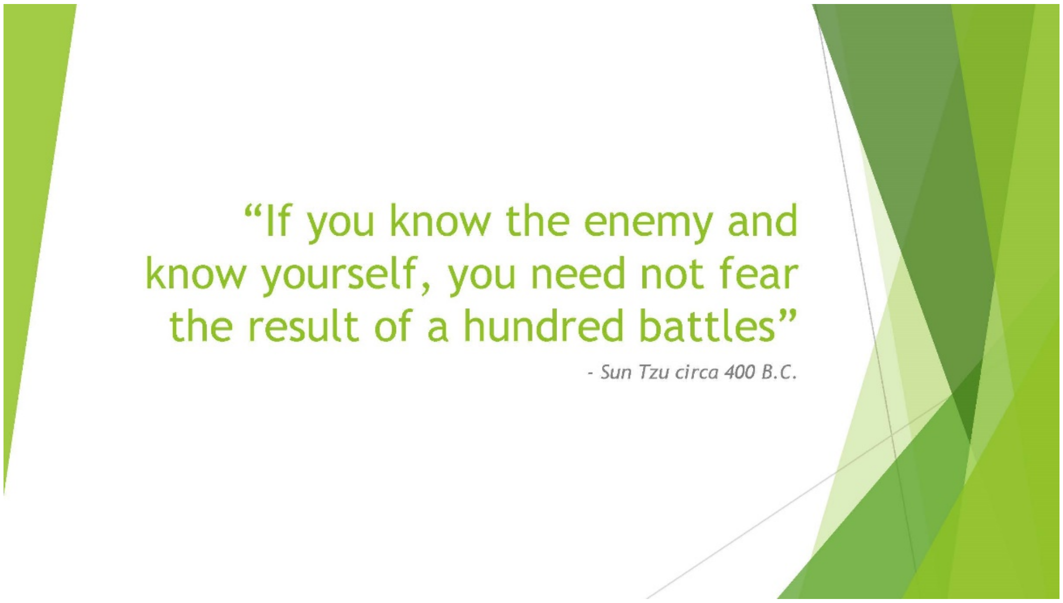
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 9



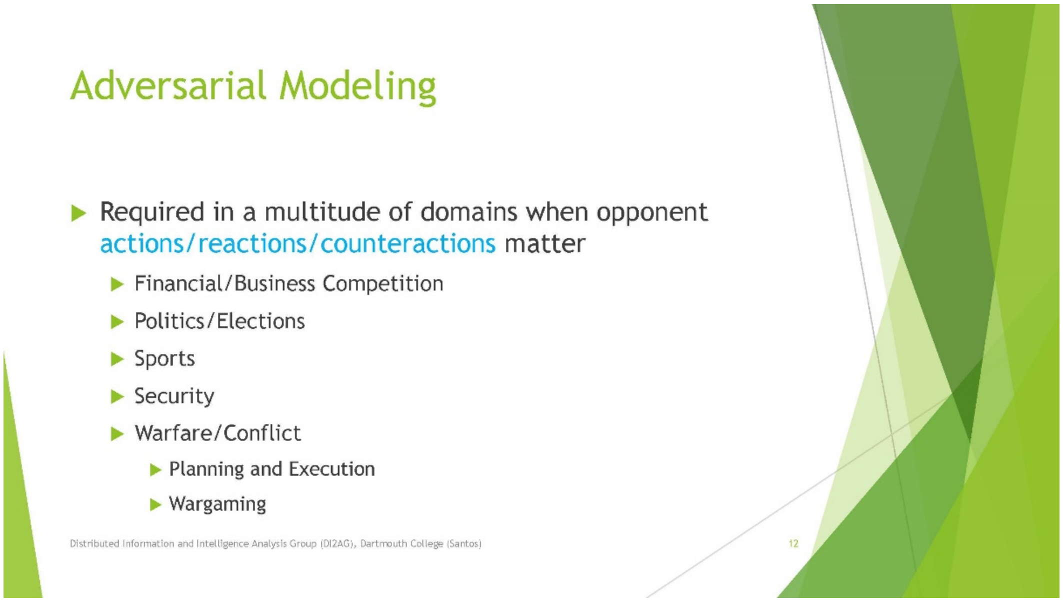
Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 10



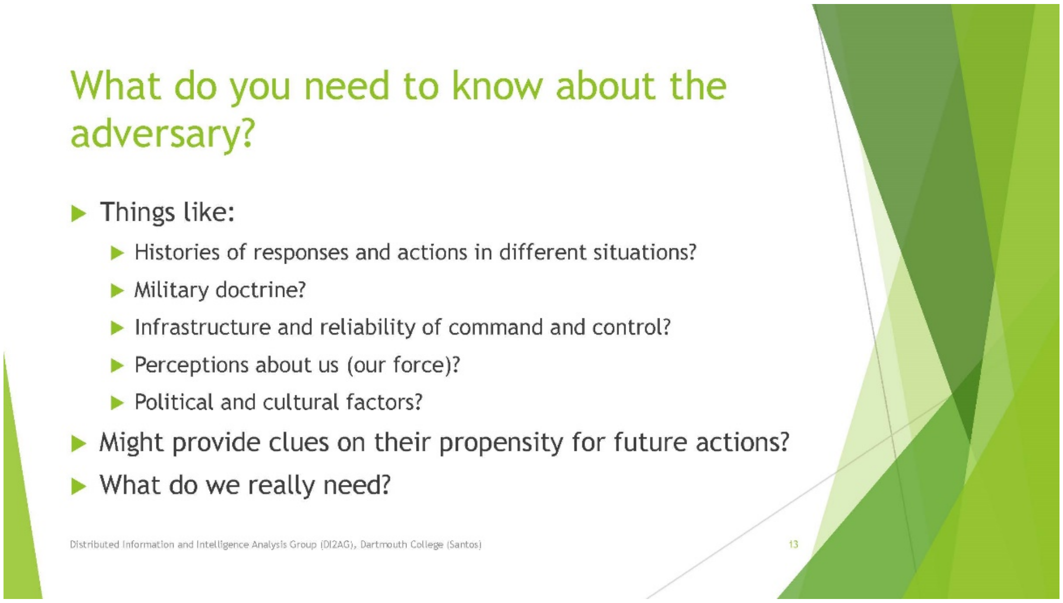
> "If you know the enemy and know yourself, you need not fear the result of a hundred battles"
>
> *- Sun Tzu circa 400 B.C.*



Adversarial Modeling

- Required in a multitude of domains when opponent actions/reactions/counteractions matter
  - Financial/Business Competition
  - Politics/Elections
  - Sports
  - Security
  - Warfare/Conflict
    - Planning and Execution
    - Wargaming

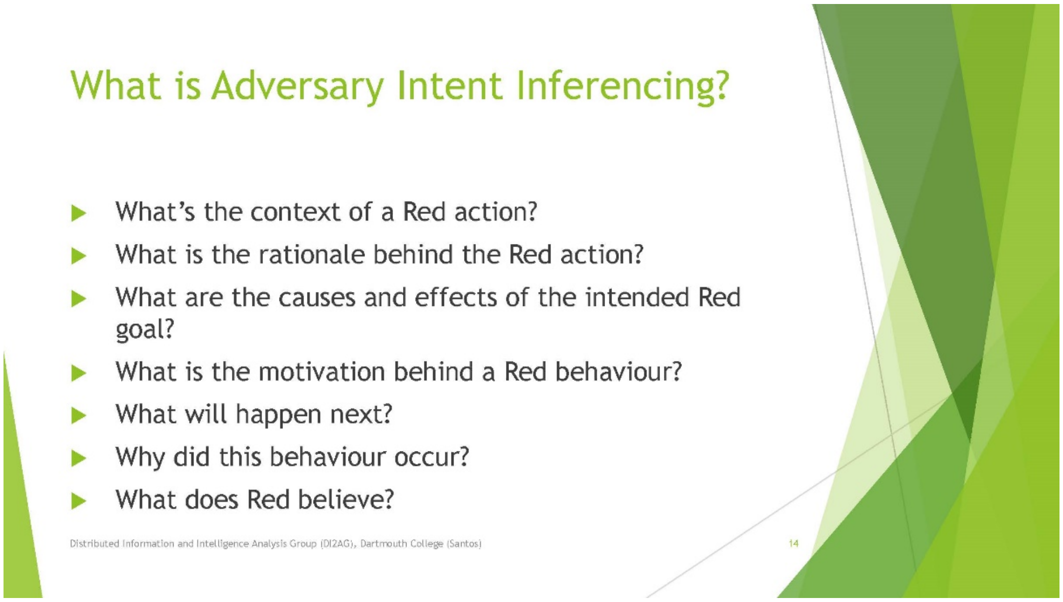
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 12



What do you need to know about the adversary?

- Things like:
  - Histories of responses and actions in different situations?
  - Military doctrine?
  - Infrastructure and reliability of command and control?
  - Perceptions about us (our force)?
  - Political and cultural factors?
- Might provide clues on their propensity for future actions?
- What do we really need?

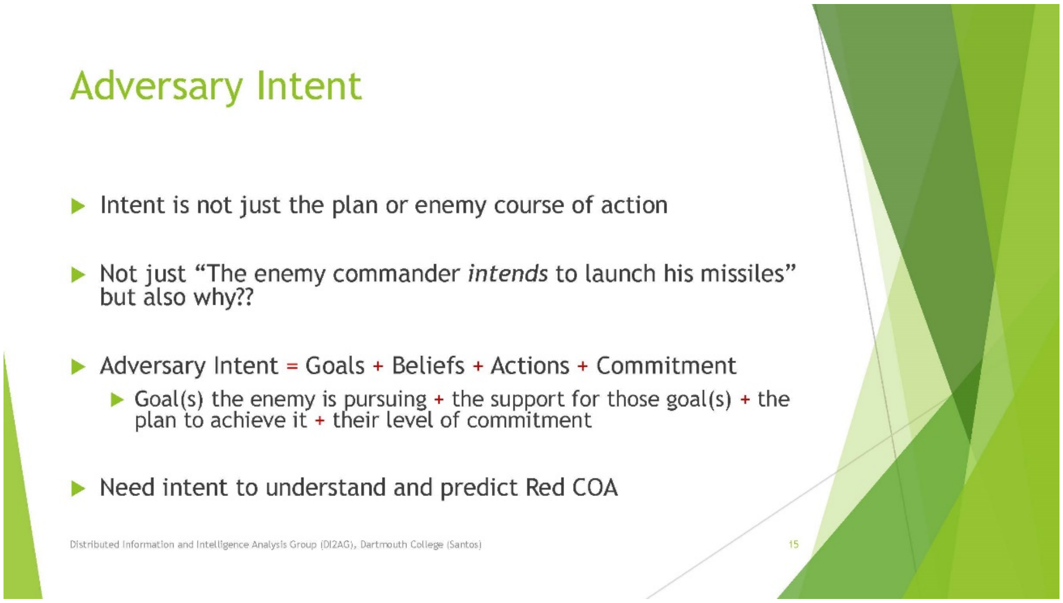
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 13



What is Adversary Intent Inferencing?

- What's the context of a Red action?
- What is the rationale behind the Red action?
- What are the causes and effects of the intended Red goal?
- What is the motivation behind a Red behaviour?
- What will happen next?
- Why did this behaviour occur?
- What does Red believe?

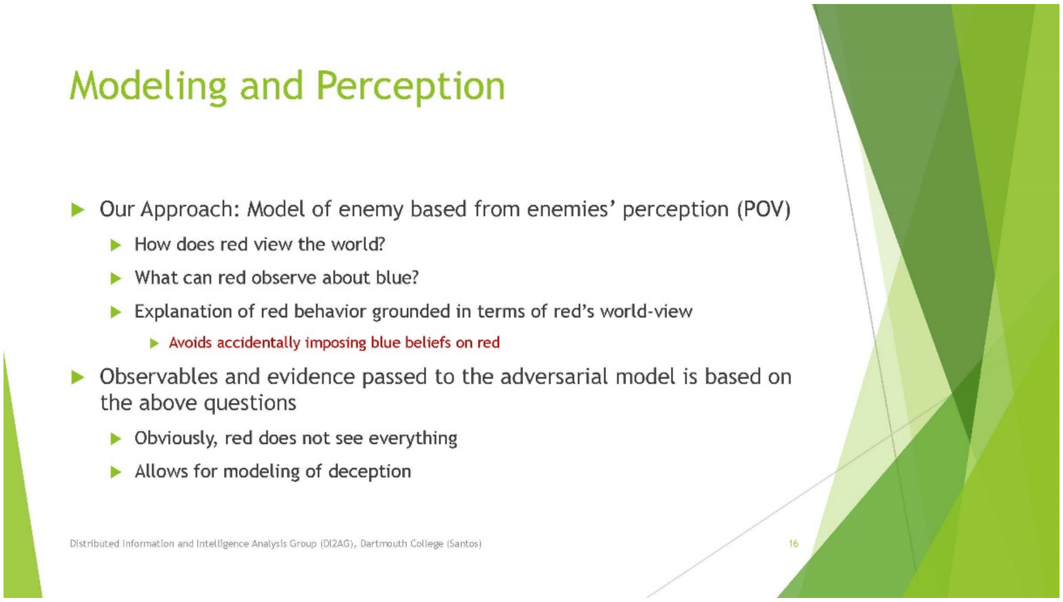
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 14



Adversary Intent

- Intent is not just the plan or enemy course of action
- Not just "The enemy commander *intends* to launch his missiles" but also why??
- Adversary Intent = Goals + Beliefs + Actions + Commitment
  - Goal(s) the enemy is pursuing + the support for those goal(s) + the plan to achieve it + their level of commitment
- Need intent to understand and predict Red COA

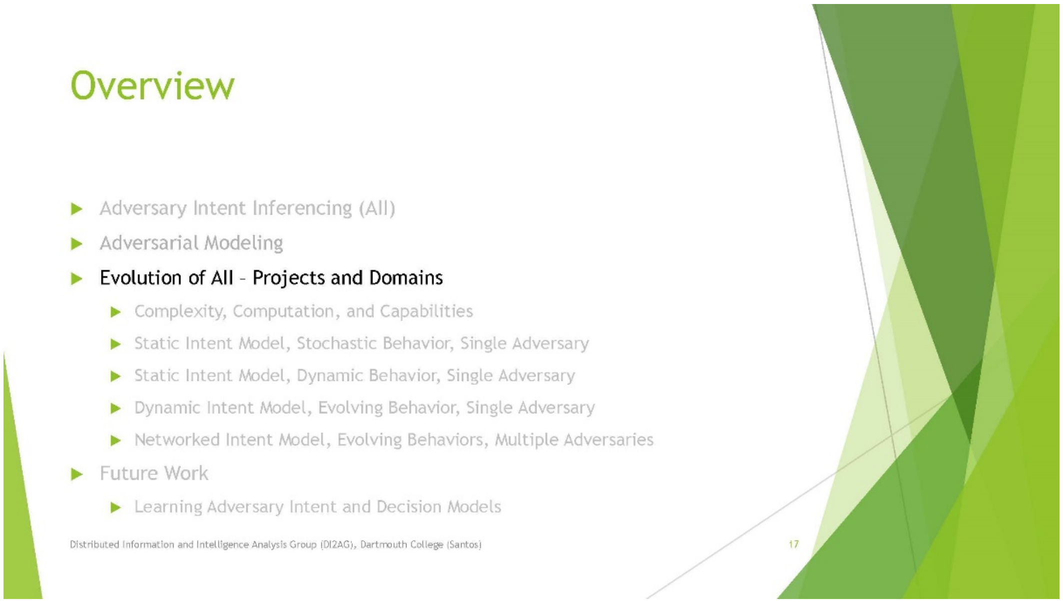
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 15



Modeling and Perception

- Our Approach: Model of enemy based from enemies' perception (POV)
  - How does red view the world?
  - What can red observe about blue?
  - Explanation of red behavior grounded in terms of red's world-view
    - Avoids accidentally imposing blue beliefs on red
- Observables and evidence passed to the adversarial model is based on the above questions
  - Obviously, red does not see everything
  - Allows for modeling of deception

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 16



Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 17

Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 18

Information Institute Research Project

Adversary Intent Inferencing for Predictive Battlespace Awareness

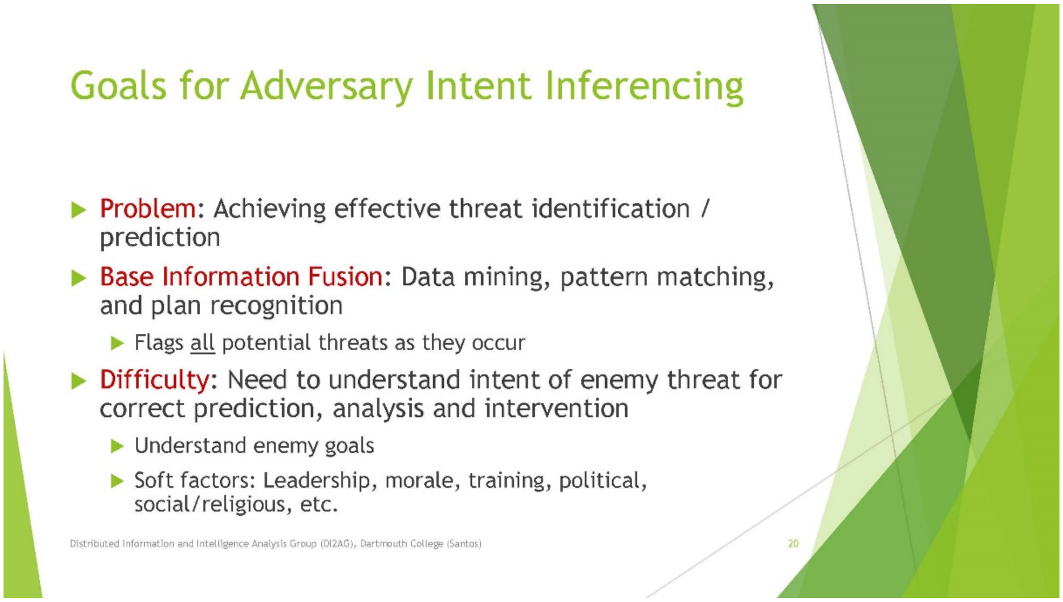
Eugene Santos Jr.
Department of Computer Science and Engineering
University of Connecticut
eugene@cse.uconn.edu
www.cse.uconn.edu/~eugene

Period of Performance: 2001 – 2004

Project Personnel

Eugene Santos Jr., UConn
Benjamin Bell, LMCO/ATL
Daniel Davenport, LMCO/ATL

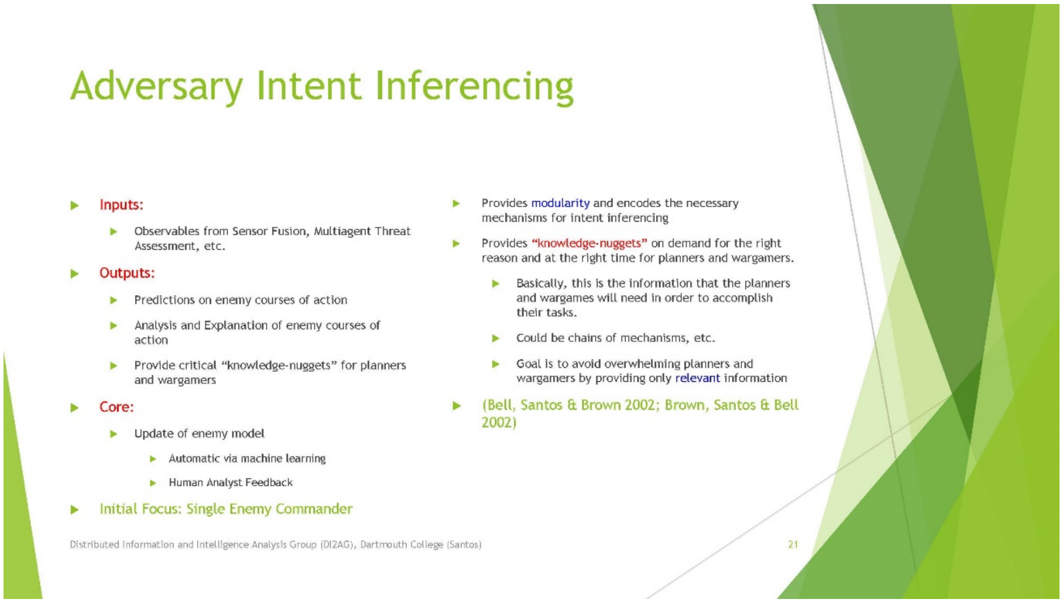
- **Bob Eggleston**, AFRL/HECA
- **2Lt Sabina Noll**, AFRL/HECA
- **Capt Scott M. Brown**, ESC



Goals for Adversary Intent Inferencing

- Problem: Achieving effective threat identification / prediction
- Base Information Fusion: Data mining, pattern matching, and plan recognition
  - Flags all potential threats as they occur
- Difficulty: Need to understand intent of enemy threat for correct prediction, analysis and intervention
  - Understand enemy goals
  - Soft factors: Leadership, morale, training, political, social/religious, etc.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 20

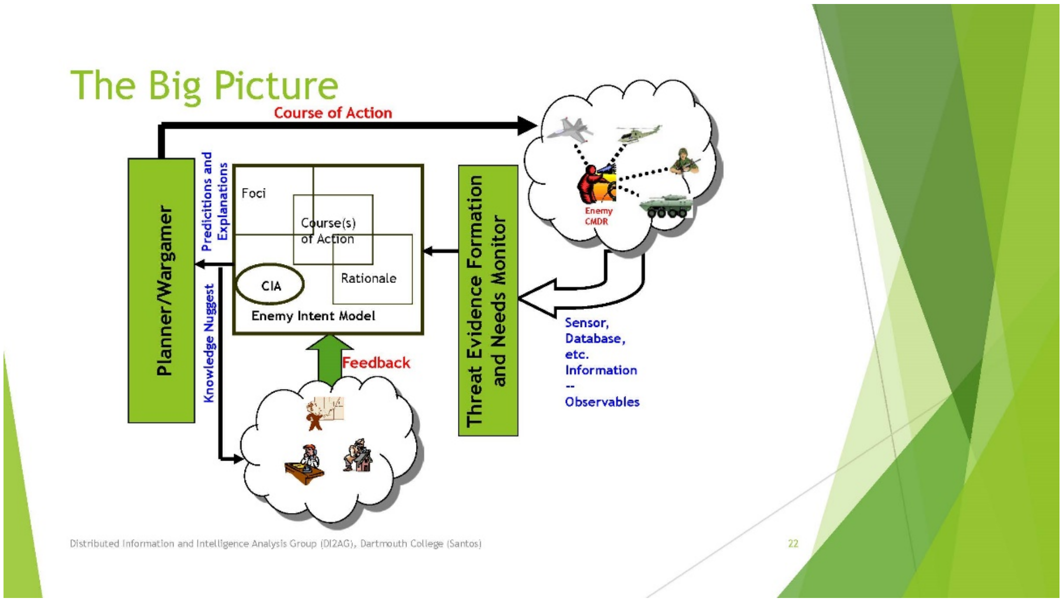


Adversary Intent Inferencing

- Inputs:
  - Observables from Sensor Fusion, Multiagent Threat Assessment, etc.
- Outputs:
  - Predictions on enemy courses of action
  - Analysis and Explanation of enemy courses of action
  - Provide critical "knowledge-nuggets" for planners and wargamers
- Core:
  - Update of enemy model
    - Automatic via machine learning
    - Human Analyst Feedback
- Initial Focus: Single Enemy Commander

- Provides modularity and encodes the necessary mechanisms for intent inferencing
- Provides "knowledge-nuggets" on demand for the right reason and at the right time for planners and wargamers.
  - Basically, this is the information that the planners and wargames will need in order to accomplish their tasks.
  - Could be chains of mechanisms, etc.
  - Goal is to avoid overwhelming planners and wargamers by providing only relevant information
- (Bell, Santos & Brown 2002; Brown, Santos & Bell 2002)

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 21



The Big Picture

Course of Action

Planner/Wargamer

Predicitions and Explanations

Knowledge Nuggest

Foci

Course(s) of Action

Rationale

CIA

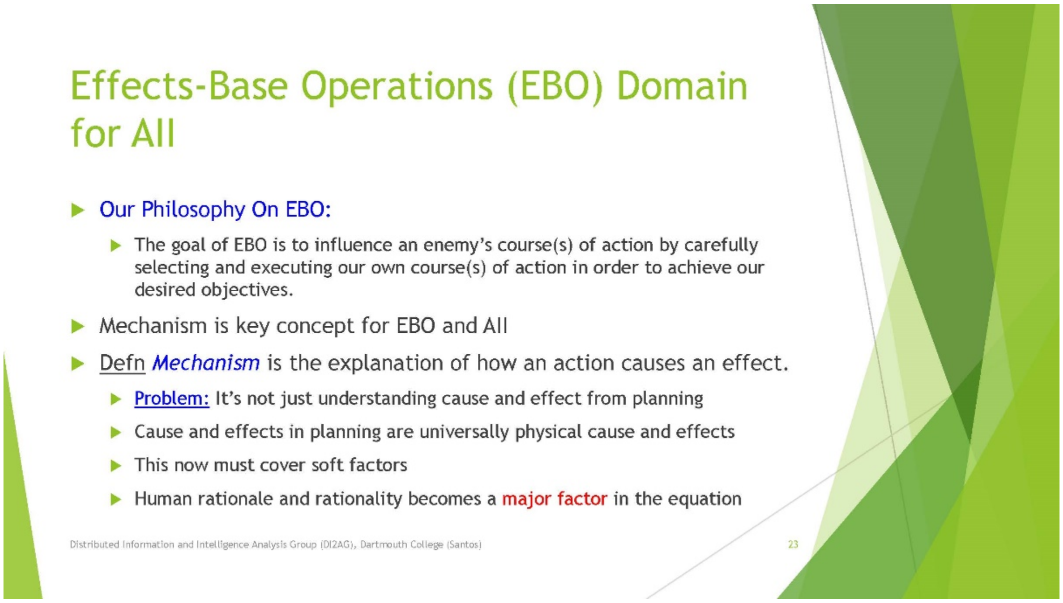
Enemy Intent Model

Feedback

Threat Evidence Formation and Needs Monitor

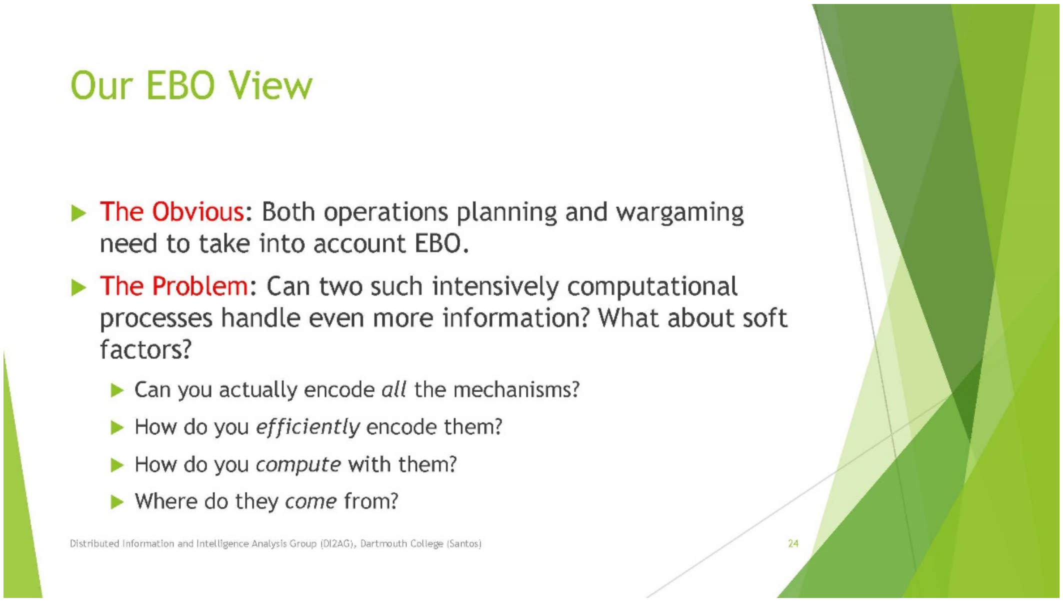
Enemy CMDR

Sensor,
Database,
etc.
Information
--
Observables



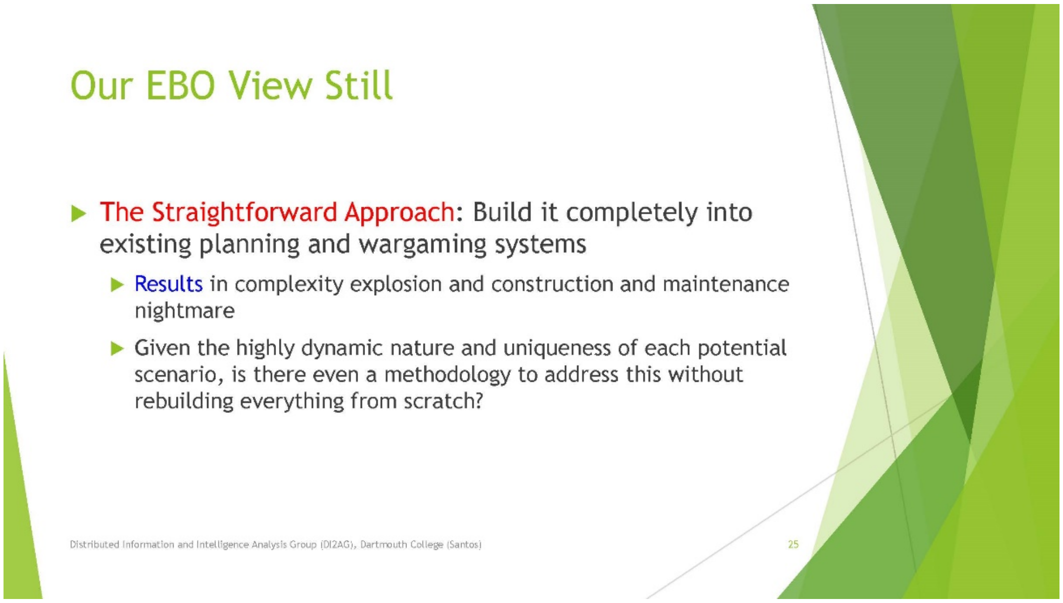
Effects-Base Operations (EBO) Domain for AII

- Our Philosophy On EBO:
  - The goal of EBO is to influence an enemy's course(s) of action by carefully selecting and executing our own course(s) of action in order to achieve our desired objectives.
- Mechanism is key concept for EBO and AII
- Defn *Mechanism* is the explanation of how an action causes an effect.
  - Problem: It's not just understanding cause and effect from planning
  - Cause and effects in planning are universally physical cause and effects
  - This now must cover soft factors
  - Human rationale and rationality becomes a major factor in the equation



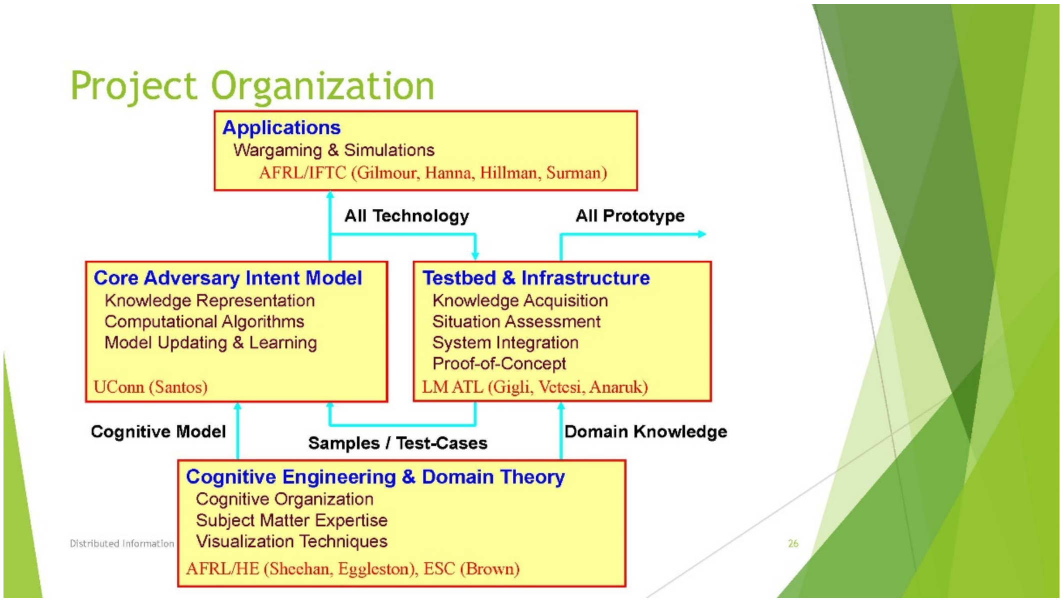
Our EBO View

- The Obvious: Both operations planning and wargaming need to take into account EBO.
- The Problem: Can two such intensively computational processes handle even more information? What about soft factors?
  - Can you actually encode *all* the mechanisms?
  - How do you *efficiently* encode them?
  - How do you *compute* with them?
  - Where do they *come* from?



Our EBO View Still

- The Straightforward Approach: Build it completely into existing planning and wargaming systems
  - Results in complexity explosion and construction and maintenance nightmare
  - Given the highly dynamic nature and uniqueness of each potential scenario, is there even a methodology to address this without rebuilding everything from scratch?



Project Organization

Applications
Wargaming & Simulations
AFRL/IFTC (Gilmour, Hanna, Hillman, Surman)

All Technology

All Prototype

Core Adversary Intent Model
Knowledge Representation
Computational Algorithms
Model Updating & Learning
UConn (Santos)

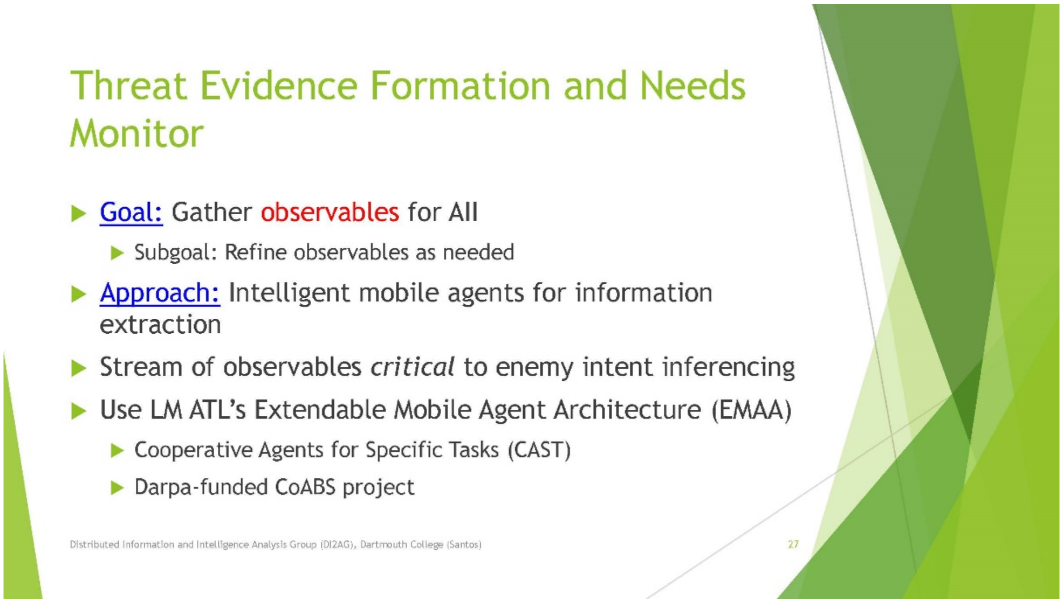
Testbed & Infrastructure
Knowledge Acquisition
Situation Assessment
System Integration
Proof-of-Concept
LM ATL (Gigli, Vetesi, Anaruk)

Cognitive Model

Samples / Test-Cases

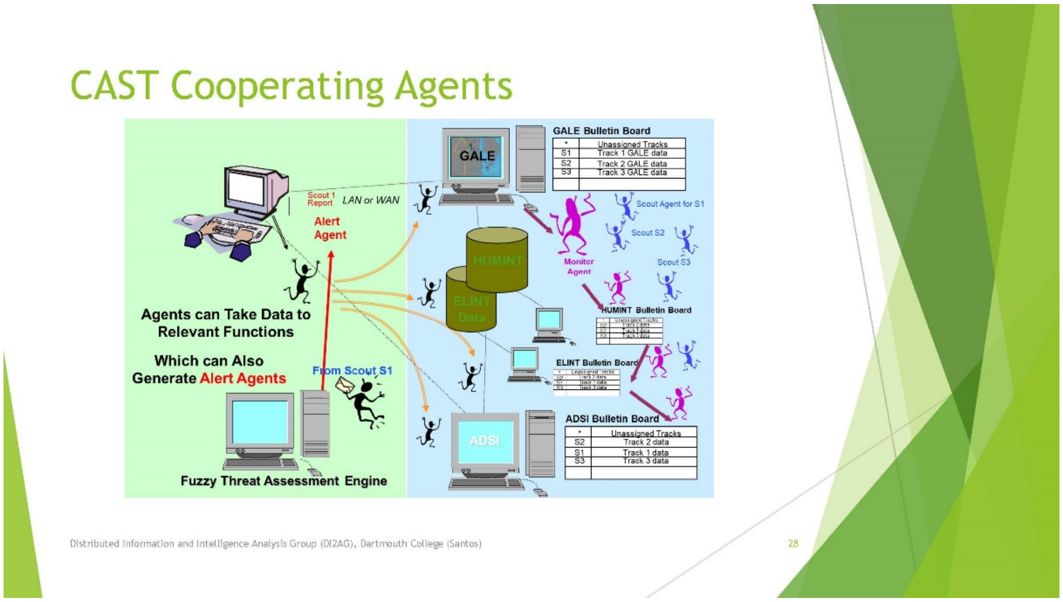
Domain Knowledge

Cognitive Engineering & Domain Theory
Cognitive Organization
Subject Matter Expertise
Visualization Techniques
AFRL/HE (Sheehan, Eggleston), ESC (Brown)



Threat Evidence Formation and Needs Monitor

- Goal: Gather observables for All
  - Subgoal: Refine observables as needed
- Approach: Intelligent mobile agents for information extraction
- Stream of observables *critical* to enemy intent inferencing
- Use LM ATL's Extendable Mobile Agent Architecture (EMAA)
  - Cooperative Agents for Specific Tasks (CAST)
  - Darpa-funded CoABS project



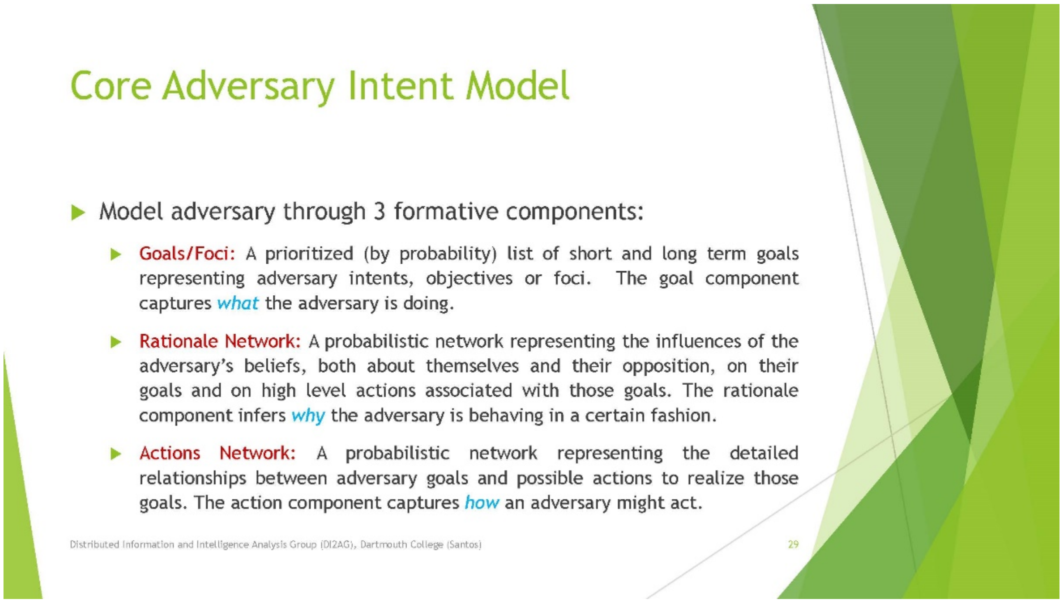
CAST Cooperating Agents

GALE Bulletin Board

*	Unassigned Tracks
S1	Track 1 GALE data
S2	Track 2 GALE data
S3	Track 3 GALE data

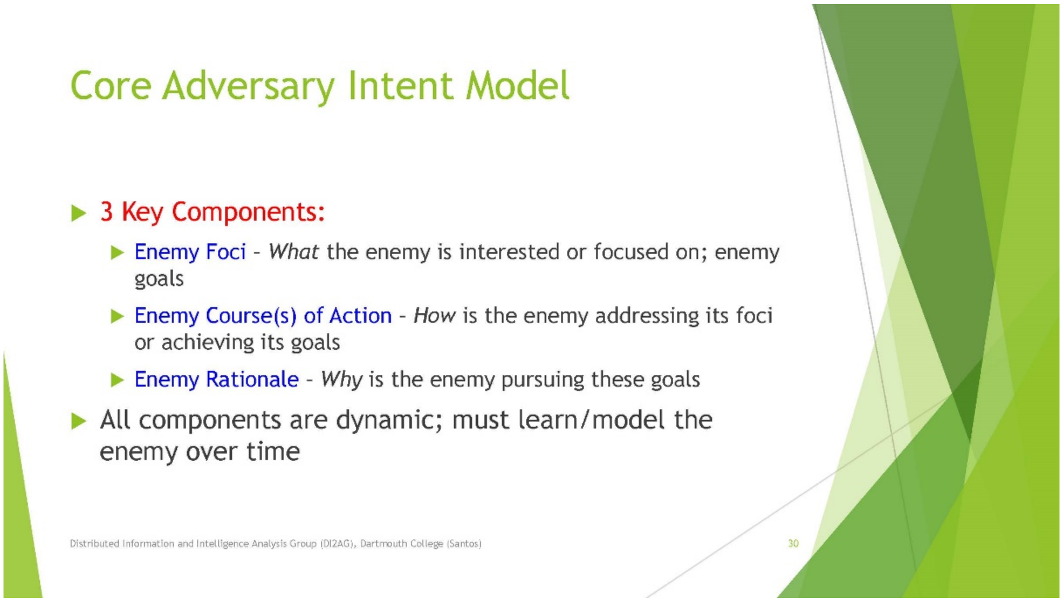
ADSI Bulletin Board

*	Unassigned Tracks
S2	Track 2 data
S1	Track 1 data
S3	Track 3 data



Core Adversary Intent Model

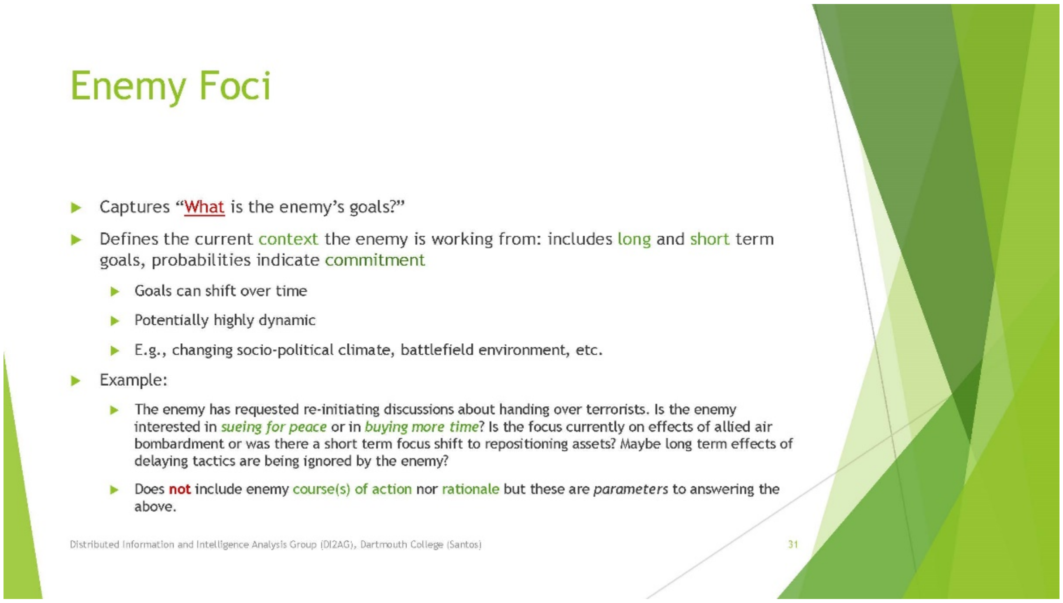
- Model adversary through 3 formative components:
  - Goals/Foci: A prioritized (by probability) list of short and long term goals representing adversary intents, objectives or foci. The goal component captures *what* the adversary is doing.
  - Rationale Network: A probabilistic network representing the influences of the adversary's beliefs, both about themselves and their opposition, on their goals and on high level actions associated with those goals. The rationale component infers *why* the adversary is behaving in a certain fashion.
  - Actions Network: A probabilistic network representing the detailed relationships between adversary goals and possible actions to realize those goals. The action component captures *how* an adversary might act.



Core Adversary Intent Model

- 3 Key Components:
  - Enemy Foci – *What* the enemy is interested or focused on; enemy goals
  - Enemy Course(s) of Action – *How* is the enemy addressing its foci or achieving its goals
  - Enemy Rationale – *Why* is the enemy pursuing these goals
- All components are dynamic; must learn/model the enemy over time

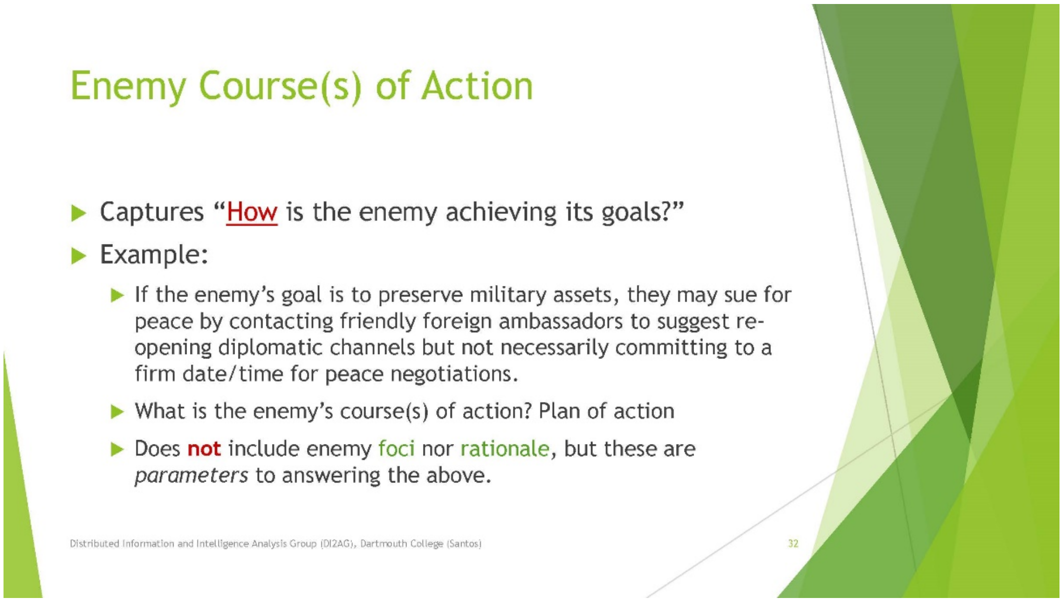
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 30



Enemy Foci

- Captures "What is the enemy's goals?"
- Defines the current context the enemy is working from: includes long and short term goals, probabilities indicate commitment
  - Goals can shift over time
  - Potentially highly dynamic
  - E.g., changing socio-political climate, battlefield environment, etc.
- Example:
  - The enemy has requested re-initiating discussions about handing over terrorists. Is the enemy interested in *sueing for peace* or in *buying more time*? Is the focus currently on effects of allied air bombardment or was there a short term focus shift to repositioning assets? Maybe long term effects of delaying tactics are being ignored by the enemy?
  - Does **not** include enemy course(s) of action nor rationale but these are *parameters* to answering the above.

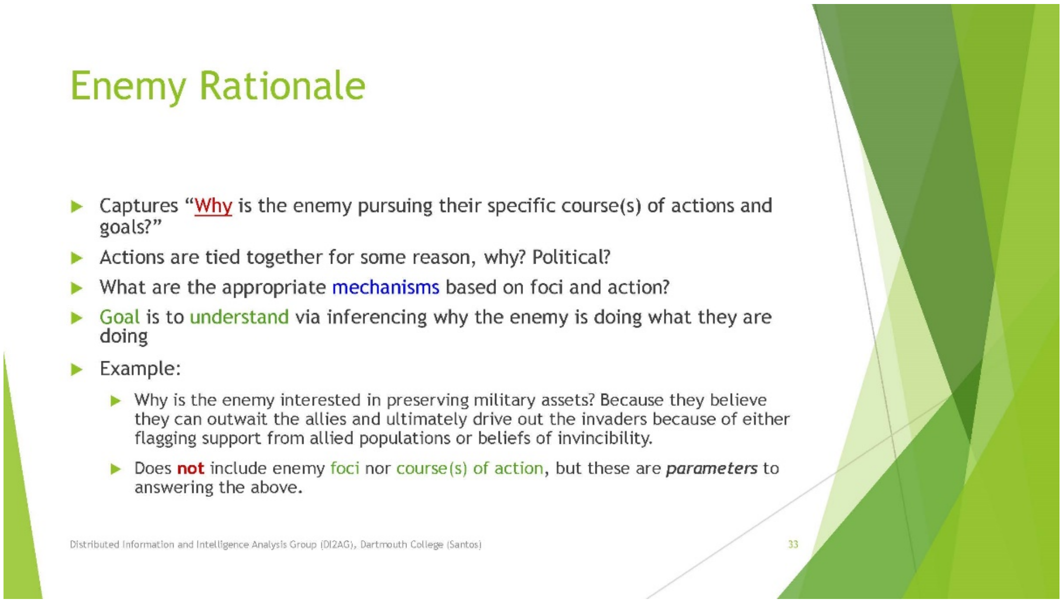
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 31



Enemy Course(s) of Action

- Captures "How is the enemy achieving its goals?"
- Example:
  - If the enemy's goal is to preserve military assets, they may sue for peace by contacting friendly foreign ambassadors to suggest re-opening diplomatic channels but not necessarily committing to a firm date/time for peace negotiations.
  - What is the enemy's course(s) of action? Plan of action
  - Does **not** include enemy foci nor rationale, but these are *parameters* to answering the above.

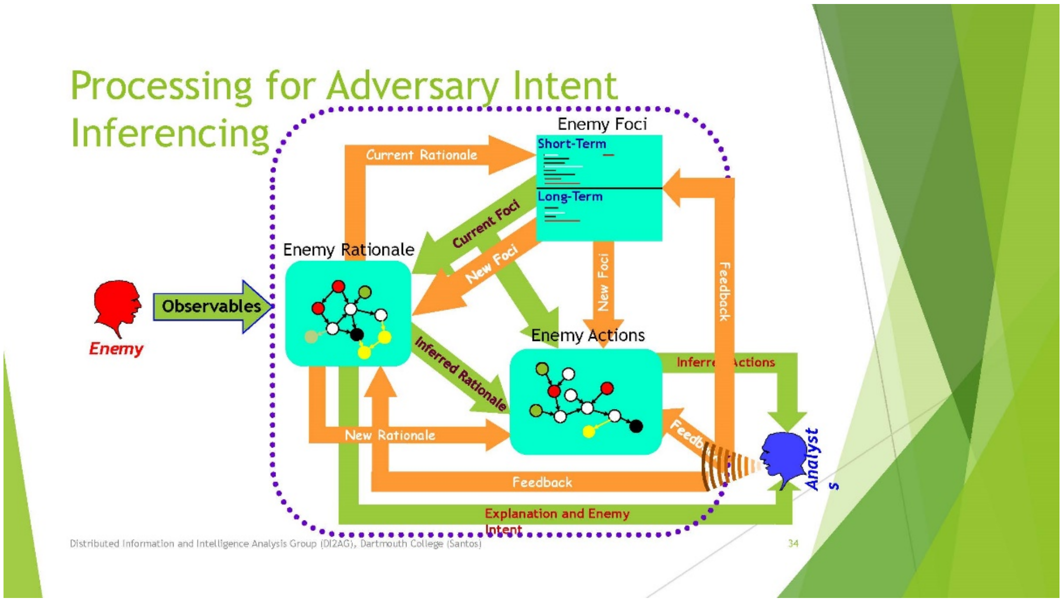
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 32



Enemy Rationale

- Captures "Why is the enemy pursuing their specific course(s) of actions and goals?"
- Actions are tied together for some reason, why? Political?
- What are the appropriate mechanisms based on foci and action?
- Goal is to understand via inferencing why the enemy is doing what they are doing
- Example:
  - Why is the enemy interested in preserving military assets? Because they believe they can outwait the allies and ultimately drive out the invaders because of either flagging support from allied populations or beliefs of invincibility.
  - Does **not** include enemy foci nor course(s) of action, but these are ***parameters*** to answering the above.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 33



Processing for Adversary Intent Inferencing

Enemy Foci

Short-Term

Long-Term

Current Rationale

Current Foci

New Foci

New Foci

Feedback

Enemy Rationale

Observables

Enemy

Enemy Actions

Inferre Actions

Inferred Rationale

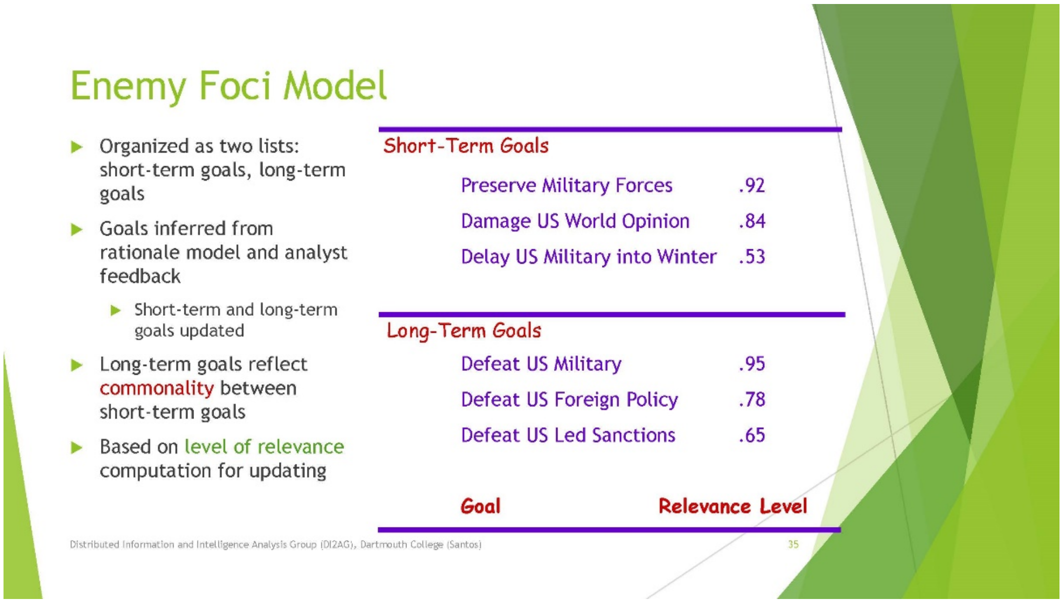
New Rationale

Feedb

Feedback

Analyst s

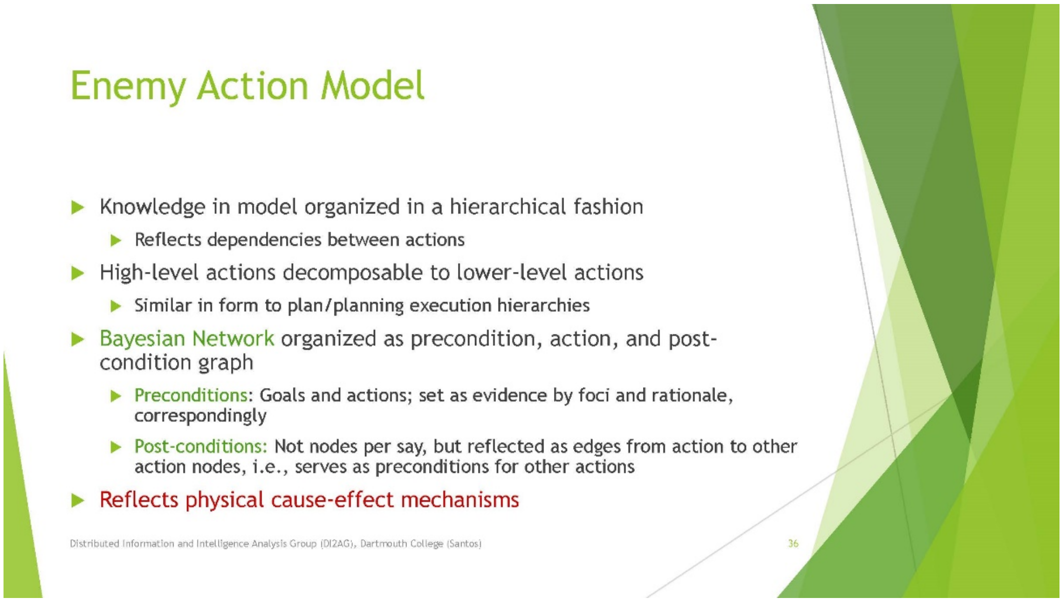
Explanation and Enemy Intent



Enemy Foci Model

- Organized as two lists: short-term goals, long-term goals
- Goals inferred from rationale model and analyst feedback
  - Short-term and long-term goals updated
- Long-term goals reflect commonality between short-term goals
- Based on level of relevance computation for updating

Goal	Relevance Level
Short-Term Goals	
Preserve Military Forces	.92
Damage US World Opinion	.84
Delay US Military into Winter	.53
Long-Term Goals	
Defeat US Military	.95
Defeat US Foreign Policy	.78
Defeat US Led Sanctions	.65



Enemy Action Model

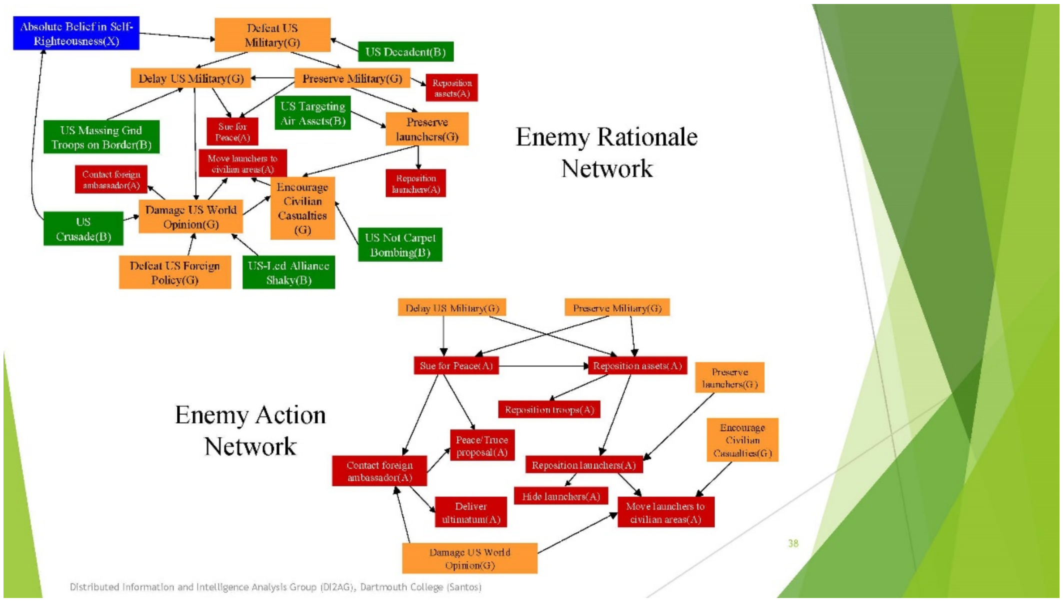
- Knowledge in model organized in a hierarchical fashion
  - Reflects dependencies between actions
- High-level actions decomposable to lower-level actions
  - Similar in form to plan/planning execution hierarchies
- Bayesian Network organized as precondition, action, and post-condition graph
  - Preconditions: Goals and actions; set as evidence by foci and rationale, correspondingly
  - Post-conditions: Not nodes per say, but reflected as edges from action to other action nodes, i.e., serves as preconditions for other actions
- Reflects physical cause-effect mechanisms

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 36

Enemy Rationale Model

- Consists of a hierarchy of goals and subgoals
  - Short-term goals are foundations for long-term goals; goals drive actions via enemy action model
- Enemy goals driven by enemy observations and perceptions of **our** actions and goals
- Bayes net reflects subgoals—goals—action organization
  - Subgoals and goals are set as evidence by enemy foci and ours and theirs observables
  - New goals are inferred here
- Mechanisms for explaining *goals* and *goal relationships* as well as *goals to our actions*

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 37

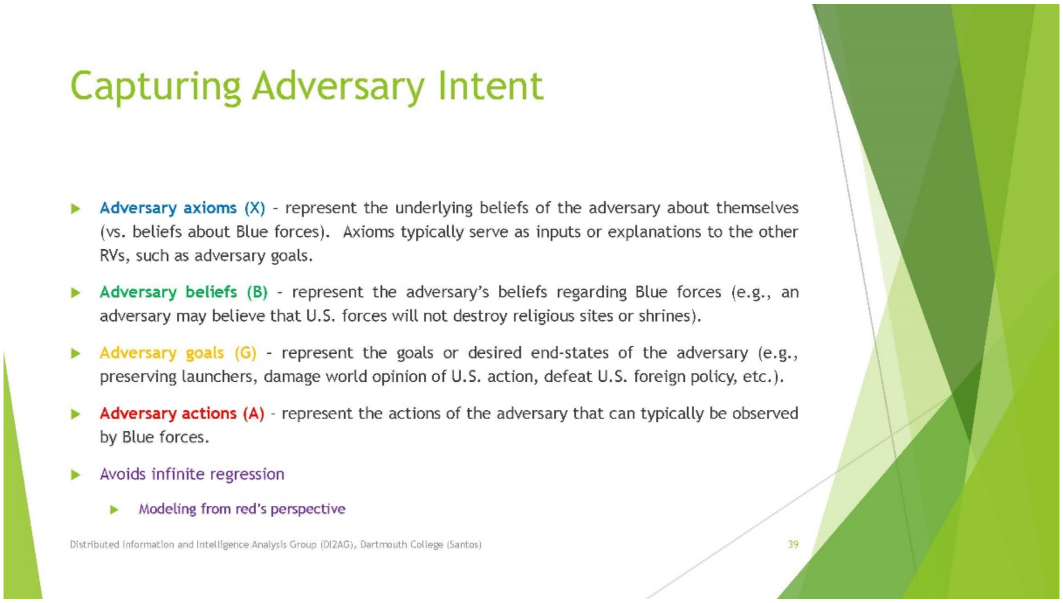

Absolute Belief in Self-Righteousness(X)
Defeat US Military(G)
US Decadent(B)
Delay US Military(G)
Preserve Military(G)
Reposition assets(A)
US Targeting Air Assets(B)
US Massing Gnd Troops on Border(B)
Sue for Peace(A)
Preserve launchers(G)
Move launchers to civilian areas(A)
Contact foreign ambassador(A)
Encourage Civilian Casualties (G)
Reposition launchers(A)
US Crusade(B)
Damage US World Opinion(G)
US Not Carpet Bombing(B)
Defeat US Foreign Policy(G)
US-Led Alliance Shaky(B)

Enemy Rationale Network

Enemy Action Network

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

38



Capturing Adversary Intent

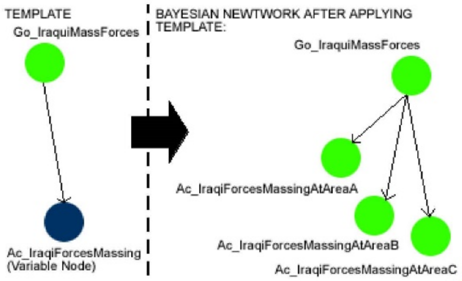
- **Adversary axioms (X)** – represent the underlying beliefs of the adversary about themselves (vs. beliefs about Blue forces). Axioms typically serve as inputs or explanations to the other RVs, such as adversary goals.
- **Adversary beliefs (B)** – represent the adversary's beliefs regarding Blue forces (e.g., an adversary may believe that U.S. forces will not destroy religious sites or shrines).
- **Adversary goals (G)** – represent the goals or desired end-states of the adversary (e.g., preserving launchers, damage world opinion of U.S. action, defeat U.S. foreign policy, etc.).
- **Adversary actions (A)** – represent the actions of the adversary that can typically be observed by Blue forces.
- Avoids infinite regression
  - Modeling from red's perspective

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

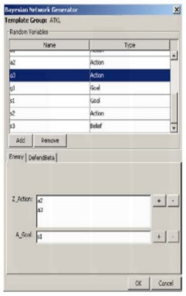
39

ATG

- All Template Generator
  - Goal: Quickly construct bayesian network components for AII
  - Uses templates to instantiate sub-networks that satisfy AII semantic structure
  - Provides a graphical drag and drop environment for quickly prototyping subnetworks
- Build and construct adversaries in a more efficient and friendly manner
- Allows for the construction of libraries of templates
- Enables the development of on-the-fly construction of adversaries and emergent adversaries


TEMPLATE
Go_IraquiMassForces
Ac_IraqiForcesMassing
(Variable Node)
BAYESIAN NEWTWORK AFTER APPLYING TEMPLATE:
Go_IraquiMassForces
Ac_IraqiForcesMassingAtAreaA
Ac_IraqiForcesMassingAtAreaB
Ac_IraqiForcesMassingAtAreaC

ATG



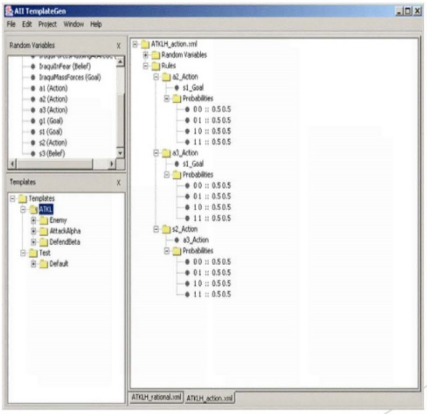
Template Group: ATKL

Random Variables

Name	Type
a2	Action
a3	Action
g1	Goal
s1	Goal
s2	Action
s3	Belief

Z_Action: a2, a3

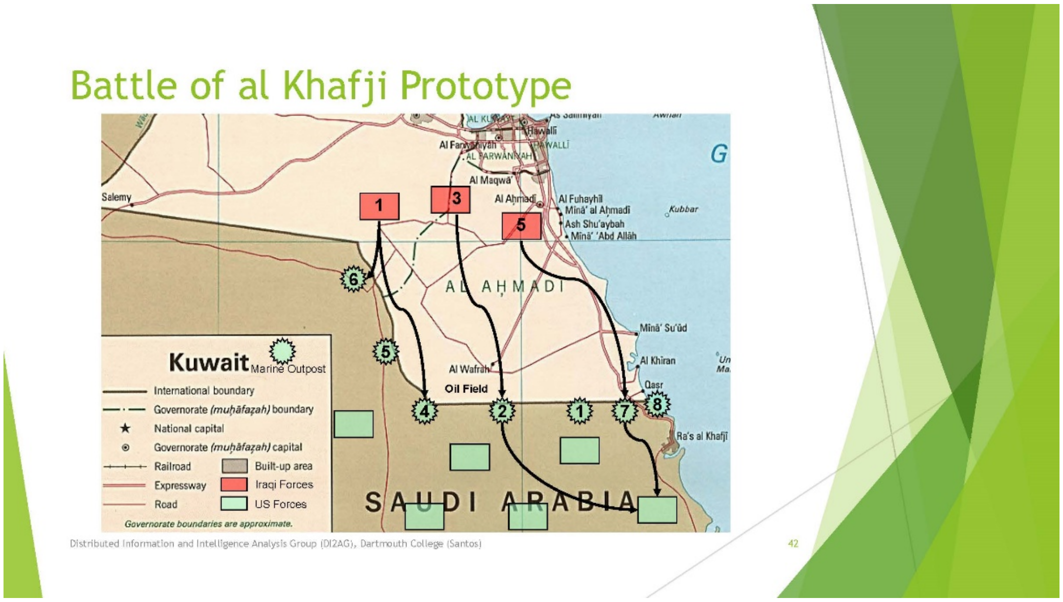
A_Goal: s1



```
IraquiInFear (Belief)
IraquiMassForces (Goal)
a1 (Action)
a2 (Action)
a3 (Action)
g1 (Goal)
s1 (Goal)
s2 (Action)
s3 (Belief)
```

```
Templates
  ATKL
    Enemy
    AttackAlpha
    DefendBeta
  Test
    Default
```

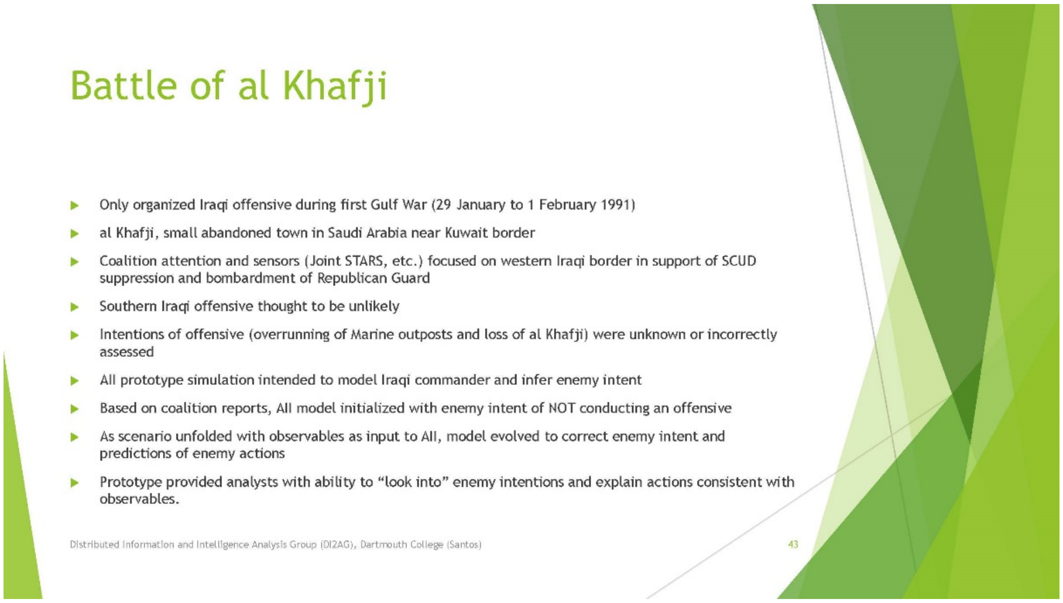
```
ATKLH_action.xml
  Random Variables
  Rules
    a2_Action
      s1_Goal
      Probabilities
        0 0 :: 0.5 0.5
        0 1 :: 0.5 0.5
        1 0 :: 0.5 0.5
        1 1 :: 0.5 0.5
    a3_Action
      s1_Goal
      Probabilities
        0 0 :: 0.5 0.5
        0 1 :: 0.5 0.5
        1 0 :: 0.5 0.5
        1 1 :: 0.5 0.5
    s2_Action
      a3_Action
      Probabilities
        0 0 :: 0.5 0.5
        0 1 :: 0.5 0.5
        1 0 :: 0.5 0.5
        1 1 :: 0.5 0.5
```



Battle of al Khafji Prototype

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

42



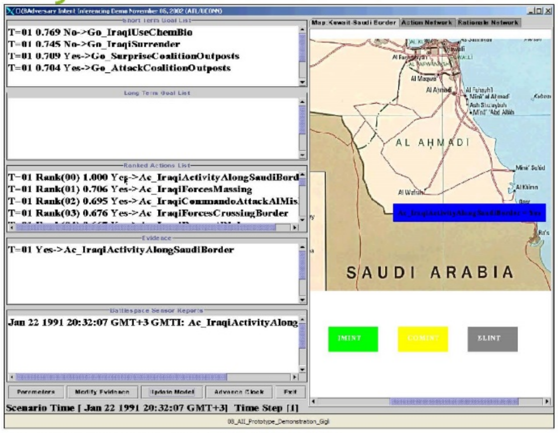
Battle of al Khafji

- Only organized Iraqi offensive during first Gulf War (29 January to 1 February 1991)
- al Khafji, small abandoned town in Saudi Arabia near Kuwait border
- Coalition attention and sensors (Joint STARS, etc.) focused on western Iraqi border in support of SCUD suppression and bombardment of Republican Guard
- Southern Iraqi offensive thought to be unlikely
- Intentions of offensive (overrunning of Marine outposts and loss of al Khafji) were unknown or incorrectly assessed
- AII prototype simulation intended to model Iraqi commander and infer enemy intent
- Based on coalition reports, AII model initialized with enemy intent of NOT conducting an offensive
- As scenario unfolded with observables as input to AII, model evolved to correct enemy intent and predictions of enemy actions
- Prototype provided analysts with ability to “look into” enemy intentions and explain actions consistent with observables.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

43

Al Khafji Simulation Screenshot



```
T=01 0.769 No->Go_IraqiUseChemBio
T=01 0.745 No->Go_IraqiSurrender
T=01 0.709 Yes->Go_SurpriseCoalitionOutposts
T=01 0.704 Yes->Go_AttackCoalitionOutposts
```

```
T=01 Rank(00) 1.000 Yes->Ac_IraqiActivityAlongSaudiBord
T=01 Rank(01) 0.706 Yes->Ac_IraqiForcesMassing
T=01 Rank(02) 0.695 Yes->Ac_IraqiCommandoAttackAlMis
T=01 Rank(03) 0.676 Yes->Ac_IraqiForcesCrossingBorder
```

```
T=01 Yes->Ac_IraqiActivityAlongSaudiBorder
```

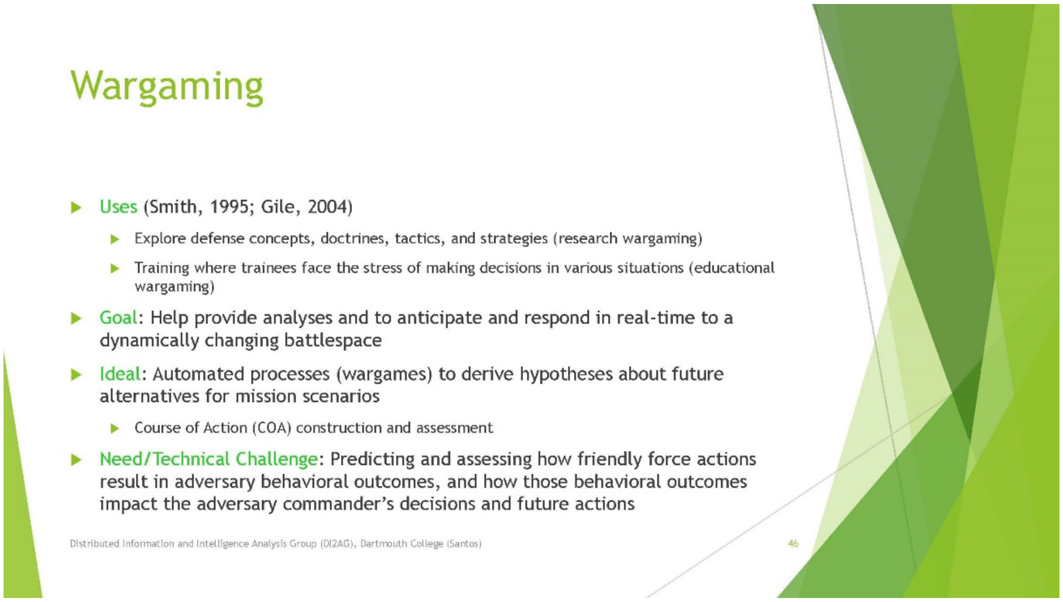
```
Jan 22 1991 20:32:07 GMT+3 GMTI: Ac_IraqiActivityAlong
```

Scenario Time [Jan 22 1991 20:32:07 GMT+3] Time Step [1]

Santos, Eugene, Jr. and Zhao, Qunhua, "Adversarial Models for Opponent Intent Inferencing," *Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind* (Eds. A. Kott and W. McEneaney), 1-22, CRC Press, 2006.

Increased Demands on the Planning Paradigm (circa 2001)

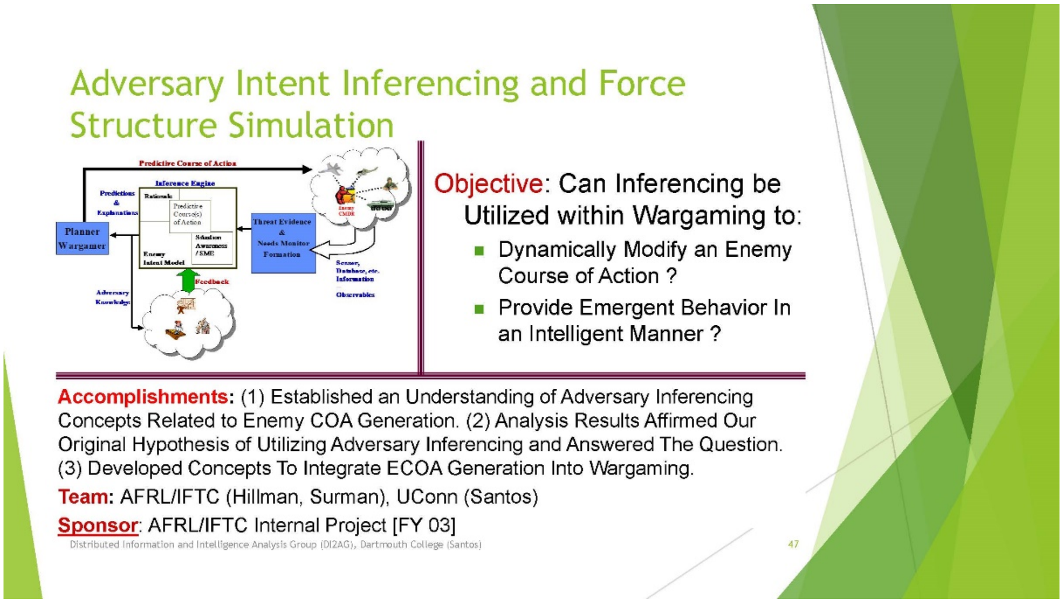
- Traditionally, Blue COAs are wargamed against the "most likely / dangerous" adversary COAs
  - Often a pre-scripted sequence of events independent of Blue actions
- Non-conventional adversaries seldom have capabilities that rival U.S. forces
  - Asymmetry of capabilities means differences in intent
- Assessment / re-assessment of friendly courses of action is limited by human capacity
- Need to model dynamic adversary behaviors that integrate with various intelligence and mission data sources (Modernized Integrated Database (MIDB), Air Operations Database (AODB), IPB Products, etc.



Wargaming

- Uses (Smith, 1995; Gile, 2004)
  - Explore defense concepts, doctrines, tactics, and strategies (research wargaming)
  - Training where trainees face the stress of making decisions in various situations (educational wargaming)
- Goal: Help provide analyses and to anticipate and respond in real-time to a dynamically changing battlespace
- Ideal: Automated processes (wargames) to derive hypotheses about future alternatives for mission scenarios
  - Course of Action (COA) construction and assessment
- Need/Technical Challenge: Predicting and assessing how friendly force actions result in adversary behavioral outcomes, and how those behavioral outcomes impact the adversary commander's decisions and future actions

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 46



Adversary Intent Inferencing and Force Structure Simulation

Predictive Course of Action
Inference Engine
Predictions & Explanations
Rationale
Predictive Course(s) of Action
Planner Wargamer
Situation Awareness / SME
Enemy Intent Model
Threat Evidence & Needs Monitor Formation
Enemy CMDR
Sensor, Database, etc. Information
Observables
Feedback
Adversary Knowledge

Objective: Can Inferencing be Utilized within Wargaming to:

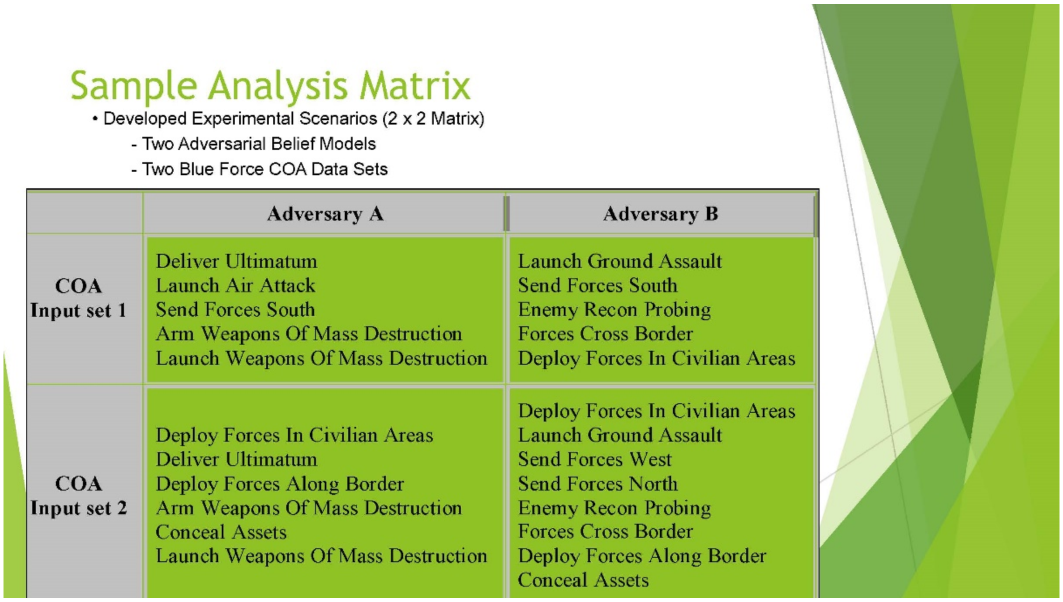
- Dynamically Modify an Enemy Course of Action ?
- Provide Emergent Behavior In an Intelligent Manner ?

Accomplishments: (1) Established an Understanding of Adversary Inferencing Concepts Related to Enemy COA Generation. (2) Analysis Results Affirmed Our Original Hypothesis of Utilizing Adversary Inferencing and Answered The Question. (3) Developed Concepts To Integrate ECOA Generation Into Wargaming.

Team: AFRL/IFTC (Hillman, Surman), UConn (Santos)

Sponsor: AFRL/IFTC Internal Project [FY 03]

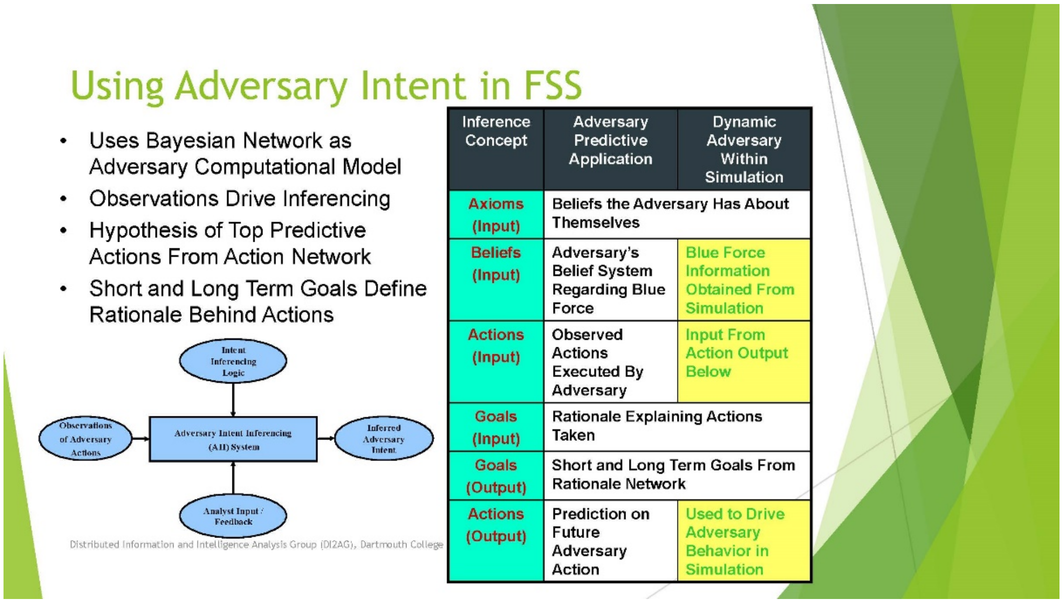
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 47



Sample Analysis Matrix

- Developed Experimental Scenarios (2 x 2 Matrix)
  - Two Adversarial Belief Models
  - Two Blue Force COA Data Sets

	Adversary A	Adversary B
COA Input set 1	Deliver Ultimatum Launch Air Attack Send Forces South Arm Weapons Of Mass Destruction Launch Weapons Of Mass Destruction	Launch Ground Assault Send Forces South Enemy Recon Probing Forces Cross Border Deploy Forces In Civilian Areas
COA Input set 2	Deploy Forces In Civilian Areas Deliver Ultimatum Deploy Forces Along Border Arm Weapons Of Mass Destruction Conceal Assets Launch Weapons Of Mass Destruction	Deploy Forces In Civilian Areas Launch Ground Assault Send Forces West Send Forces North Enemy Recon Probing Forces Cross Border Deploy Forces Along Border Conceal Assets



Using Adversary Intent in FSS

- Uses Bayesian Network as Adversary Computational Model
- Observations Drive Inferencing
- Hypothesis of Top Predictive Actions From Action Network
- Short and Long Term Goals Define Rationale Behind Actions

Intent Inferencing Logic
Observations of Adversary Actions
Adversary Intent Inferencing (AII) System
Inferred Adversary Intent
Analyst Input / Feedback

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College

Inference Concept	Adversary Predictive Application	Dynamic Adversary Within Simulation
Axioms (Input)	Beliefs the Adversary Has About Themselves	
Beliefs (Input)	Adversary's Belief System Regarding Blue Force	Blue Force Information Obtained From Simulation
Actions (Input)	Observed Actions Executed By Adversary	Input From Action Output Below
Goals (Input)	Rationale Explaining Actions Taken	
Goals (Output)	Short and Long Term Goals From Rationale Network	
Actions (Output)	Prediction on Future Adversary Action	Used to Drive Adversary Behavior in Simulation

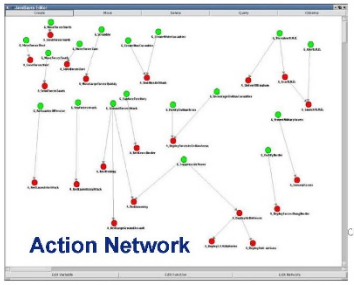
Bayesian Network Complex

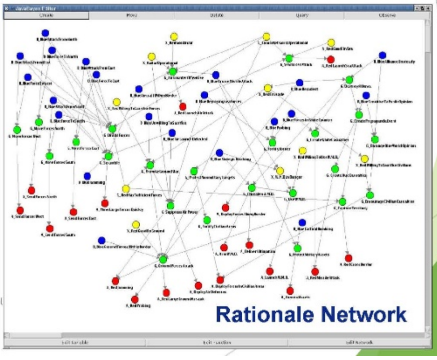
Graphical Tool Required for Static AII

1) Edit Bayesian Net 2) Visualize & Comprehend Model

- Obtained JAVA Bayesian Editor (CMU)
- JAVA Modified To Annotate Bayesian Nodes For Inferencing Concepts
- Format Converter Integrated (E. Santos)

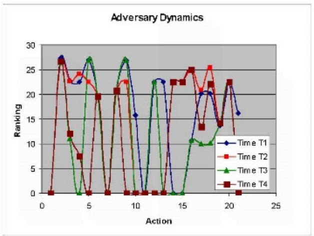
🟨 AXIOM 🟦 BELIEF 🟩 GOAL 🟥 ACTION



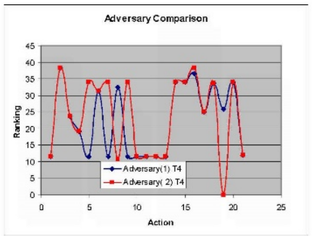


Adversary Dynamics

Dynamics Of Adversary(1) Across Time Steps



Comparison Of Adversary(1) and Adversary(2) at Time Step 4

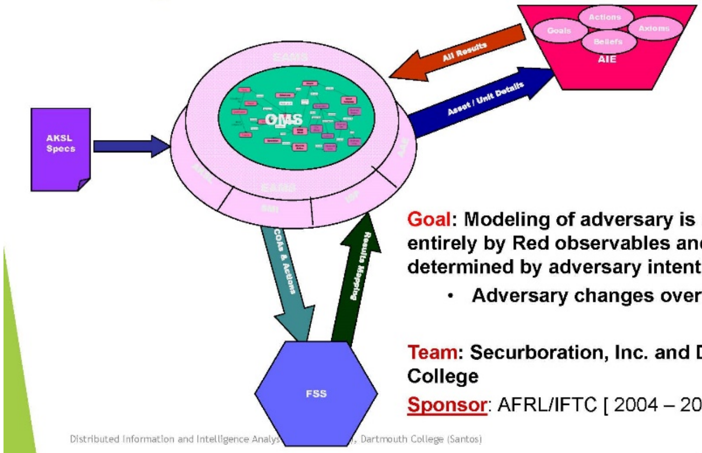


Surman, Joshua, Hillman, Robert, and Santos, Eugene, Jr., "Adversarial Inferencing for Generating Dynamic Adversary Behavior," *Proceedings of the SPIE: 17th Annual International Symposium on Aerospace/Defense Sensing and Controls: AeroSense 2003*, Vol. 5091, 194-201, Orlando, FL, 2003.

Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models

Emergent Adversarial Modeling System (EAMS)

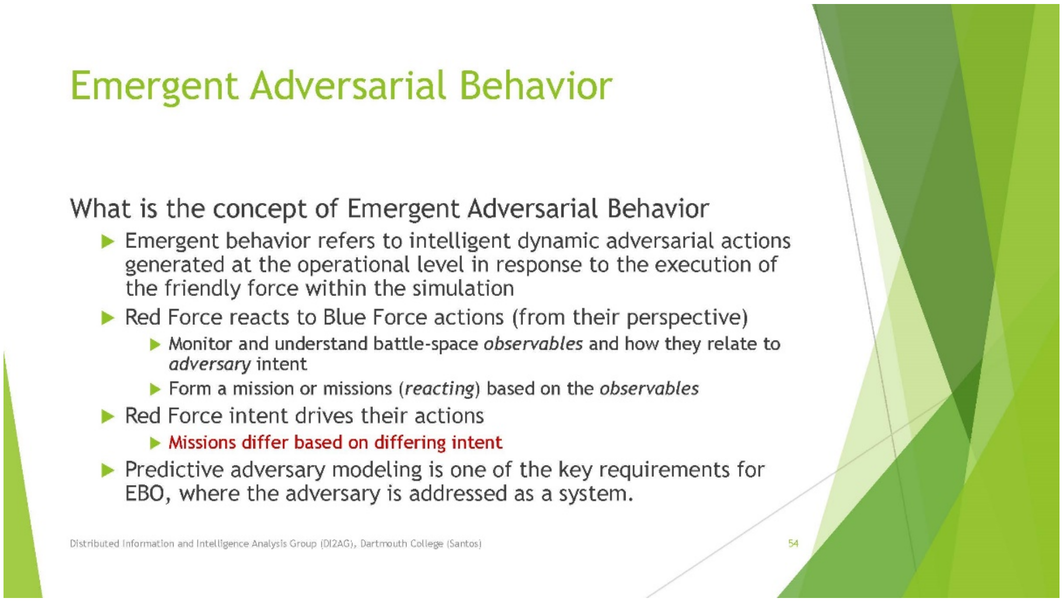


Goal: Modeling of adversary is driven entirely by Red observables and actions determined by adversary intent.

- Adversary changes over time.

Team: Securboration, Inc. and Dartmouth College

Sponsor: AFRL/IFTC [2004 – 2007]

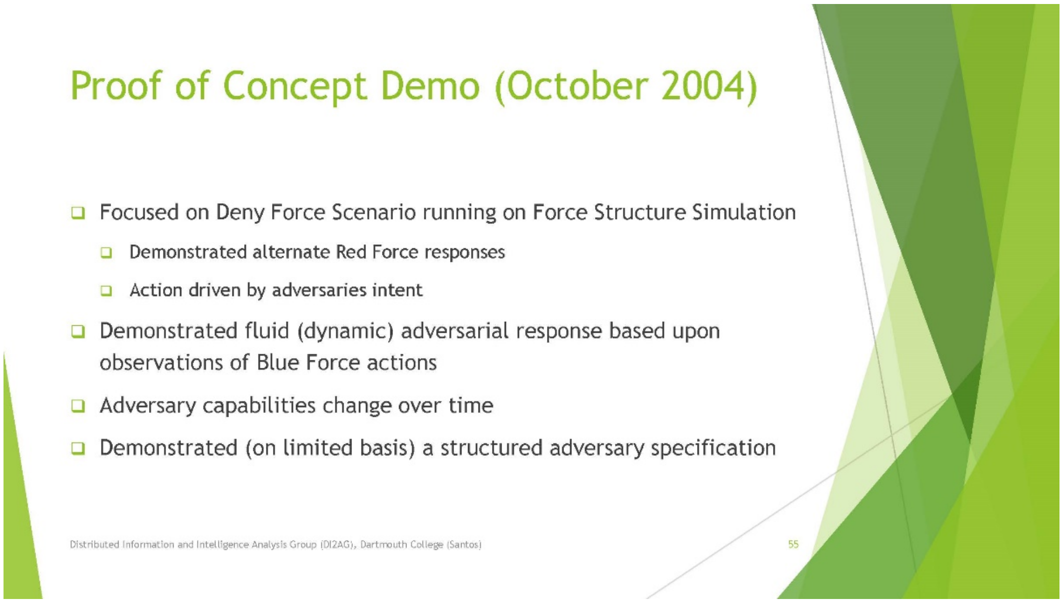


Emergent Adversarial Behavior

What is the concept of Emergent Adversarial Behavior

- ▶ Emergent behavior refers to intelligent dynamic adversarial actions generated at the operational level in response to the execution of the friendly force within the simulation
- ▶ Red Force reacts to Blue Force actions (from their perspective)
  - ▶ Monitor and understand battle-space *observables* and how they relate to *adversary* intent
  - ▶ Form a mission or missions (*reacting*) based on the *observables*
- ▶ Red Force intent drives their actions
  - ▶ Missions differ based on differing intent
- ▶ Predictive adversary modeling is one of the key requirements for EBO, where the adversary is addressed as a system.

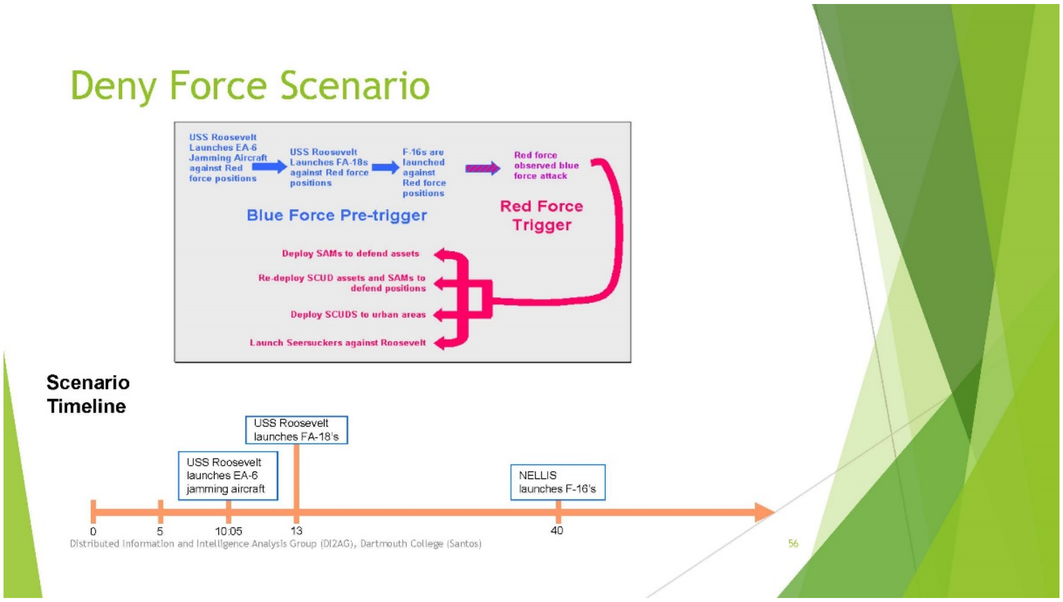
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 54



Proof of Concept Demo (October 2004)

- ❑ Focused on Deny Force Scenario running on Force Structure Simulation
  - ❑ Demonstrated alternate Red Force responses
  - ❑ Action driven by adversaries intent
- ❑ Demonstrated fluid (dynamic) adversarial response based upon observations of Blue Force actions
- ❑ Adversary capabilities change over time
- ❑ Demonstrated (on limited basis) a structured adversary specification

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 55



Deny Force Scenario

USS Roosevelt Launches EA-6 Jamming Aircraft against Red force positions

USS Roosevelt Launches FA-18s against Red force positions

F-16s are launched against Red force positions

Red force observed blue force attack

Blue Force Pre-trigger

Red Force Trigger

Deploy SAMs to defend assets

Re-deploy SCUD assets and SAMs to defend positions

Deploy SCUDS to urban areas

Launch Seersuckers against Roosevelt

Scenario Timeline

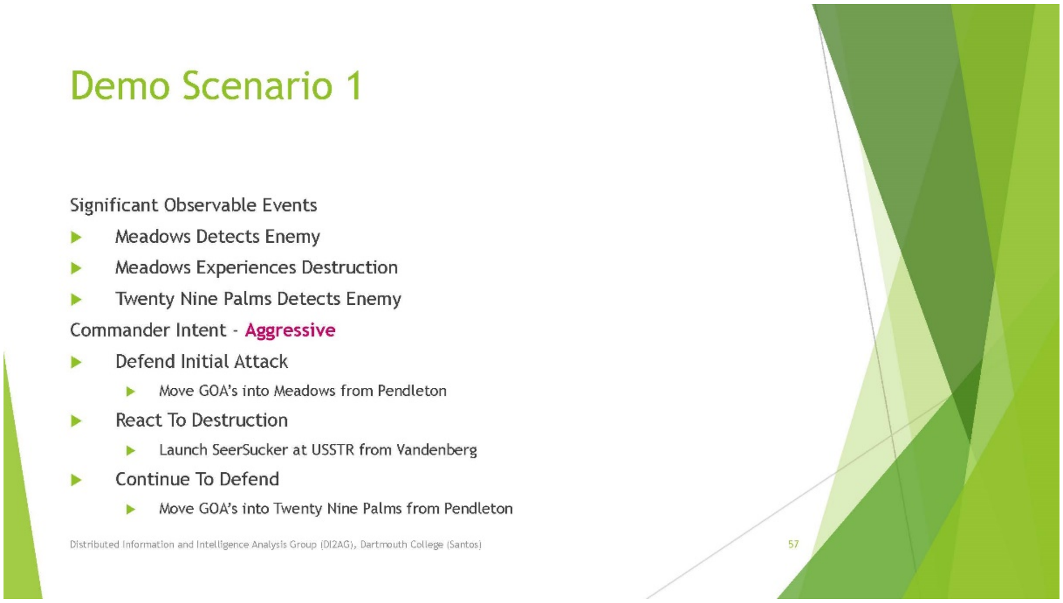
USS Roosevelt launches EA-6 jamming aircraft

USS Roosevelt launches FA-18's

NELLIS launches F-16's

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

56



Demo Scenario 1

Significant Observable Events

- Meadows Detects Enemy
- Meadows Experiences Destruction
- Twenty Nine Palms Detects Enemy

Commander Intent - **Aggressive**

- Defend Initial Attack
  - Move GOA's into Meadows from Pendleton
- React To Destruction
  - Launch SeerSucker at USSTR from Vandenberg
- Continue To Defend
  - Move GOA's into Twenty Nine Palms from Pendleton

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

57

Demo Scenario 2

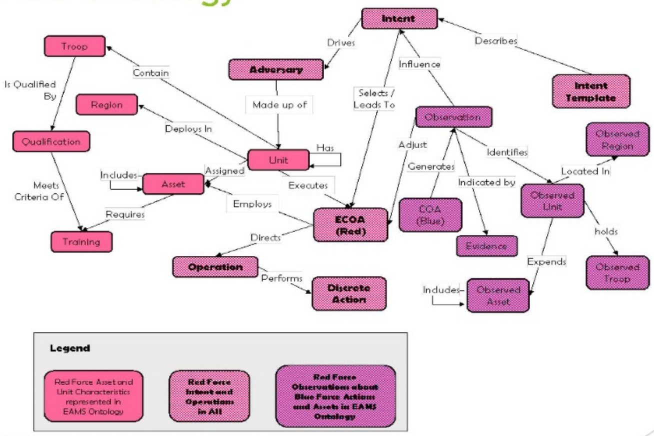
Significant Observable Events

- Meadows Detects Enemy
- Meadows Experiences Destruction
- Twenty Nine Palms Detects Enemy

Commander Intent - **Passive**

- Defend Initial Attack
  - Move GOA's into Meadows from Pendleton
- Continue To Defend
  - Move GOA's into Twenty Nine Palms from Pendleton
- Defend With Authority
  - Operate All SA-2's

EAMS Ontology



Intent
Drives
Describes
Troop
Adversary
Influence
Contain
Intent Template
Is Qualified By
Selects / Leads To
Region
Made up of
Observation
Deploys In
Qualification
Observed Region
Has
Unit
Adjust
Identifies
Generates
Located In
Includes
Assigned
Asset
Executes
Indicated by
Observed Unit
Meets Criteria Of
Employs
COA (Blue)
Requires
ECOA (Red)
holds
Evidence
Directs
Training
Expends
Operation
Observed Troop
Performs
Discrete Action
Includes
Observed Asset

Legend

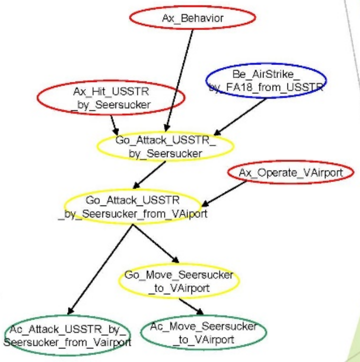
Red Force Asset and Unit Characteristics represented in EAMS Ontology

Red Force Intent and Operations in AII

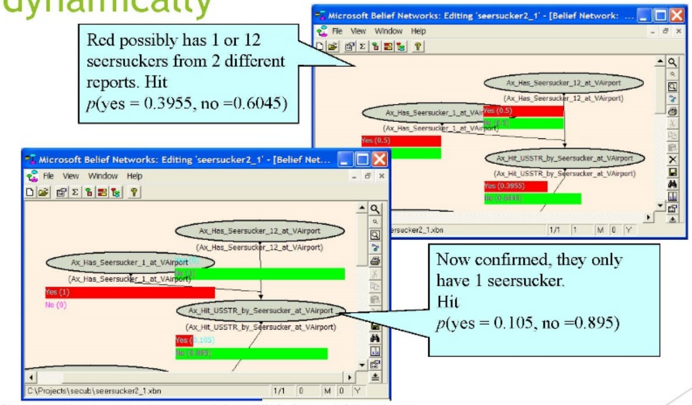
Red Force Observations about Blue Force Actions and Assets in EAMS Ontology

An example of BKF generation

- According to the scenarios, there is a goal to attack USS TR with sunburn, which is a new asset not in the working network.
  - Assets include: USS TR, sunburn, VAirport, …
- Search in the library retrieves one fragment (shown next)
  - Include, USS TR, VAirport, seersucker.
  - Also the goals, axioms, and beliefs are very similar


Ax_Behavior
Ax_Hit_USSTR _by_Seersucker
Be_AirStrike_ by_FA18_from_USSTR
Go_Attack_USSTR_ by_Seersucker
Ax_Operate_VAirport
Go_Attack_USSTR _by_Seersucker_from_VAirport
Go_Move_Seersucker _to_VAirport
Ac_Attack_USSTR_by_ Seersucker_from_Vairport
Ac_Move_Seersucker _to_VAirport

Represent numbers of assets dynamically


Microsoft Belief Networks: Editing 'seersucker2_1' - [Belief Network: …
File View Window Help
Ax_Has_Seersucker_12_at_VAirport
(Ax_Has_Seersucker_12_at_VAirport)
Yes (0.5)
No (0.5)
Ax_Has_Seersucker_1_at_VAirport
(Ax_Has_Seersucker_1_at_VAirport)
Yes (0.5)
Ax_Hit_USSTR_by_Seersucker_at_VAirport
(Ax_Hit_USSTR_by_Seersucker_at_VAirport)
Yes (0.3955)
No (0.6045)
ersucker2_1.xbn 1/1 1 M 0 Y
Microsoft Belief Networks: Editing 'seersucker2_1' - [Belief Net…
File View Window Help
Ax_Has_Seersucker_12_at_VAirport
(Ax_Has_Seersucker_12_at_VAirport)
Yes (0)
No (1)
Ax_Has_Seersucker_1_at_VAirport
(Ax_Has_Seersucker_1_at_VAirport)
Yes (1)
No (0)
Ax_Hit_USSTR_by_Seersucker_at_VAirport
(Ax_Hit_USSTR_by_Seersucker_at_VAirport)
Yes (0.105)
No (0.895)
C:\Projects\secub\seersucker2_1.xbn 1/1 0 M 0 Y

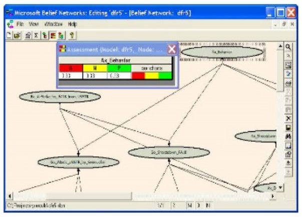
Red possibly has 1 or 12 seersuckers from 2 different reports. Hit
p(yes = 0.3955, no =0.6045)

Now confirmed, they only have 1 seersucker.
Hit
p(yes = 0.105, no =0.895)

Behavior and Affects

Ax_Behavior represents a soft factor of red (commander).

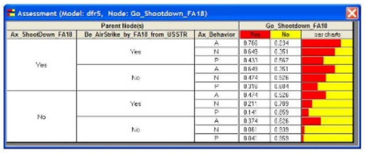
Three states:
Aggressive,
Neutral,
Passive



Microsoft Belief Networks: Editing 'dfr5' - [Belief Network: dfr5]

Assessment (Model: dfr5, Node: ...

Ax_Behavior			
A	N	P	bar charts
333	333	0.33	🟥🟨🟩



Assessment (Model: dfr5, Node: Go_Shootdown_FA18)

Parent Node(s)			Go_Shootdown_FA18		
Ax_ShootDown_FA18	Be_AirStrike_by_FA18_from_USSTR	Ax_Behavior	Yes	No	bar charts
Yes	Yes	A	0.766	0.234	
		N	0.649	0.351	
		P	0.433	0.567	
	No	A	0.649	0.351	
		N	0.474	0.526	
		P	0.316	0.684	
No	Yes	A	0.474	0.526	
		N	0.211	0.789	
		P	0.141	0.859	
	No	A	0.374	0.626	
		N	0.061	0.939	
		P	0.041	0.959	

Assume the probability for the neutral states (N) is p_n,
The Probability for aggressive states (A) is: $p_n + 0.33 * (1.0 - p_n)$
The Probability for passive states (P) is: $(1- 0.33) * p_n$

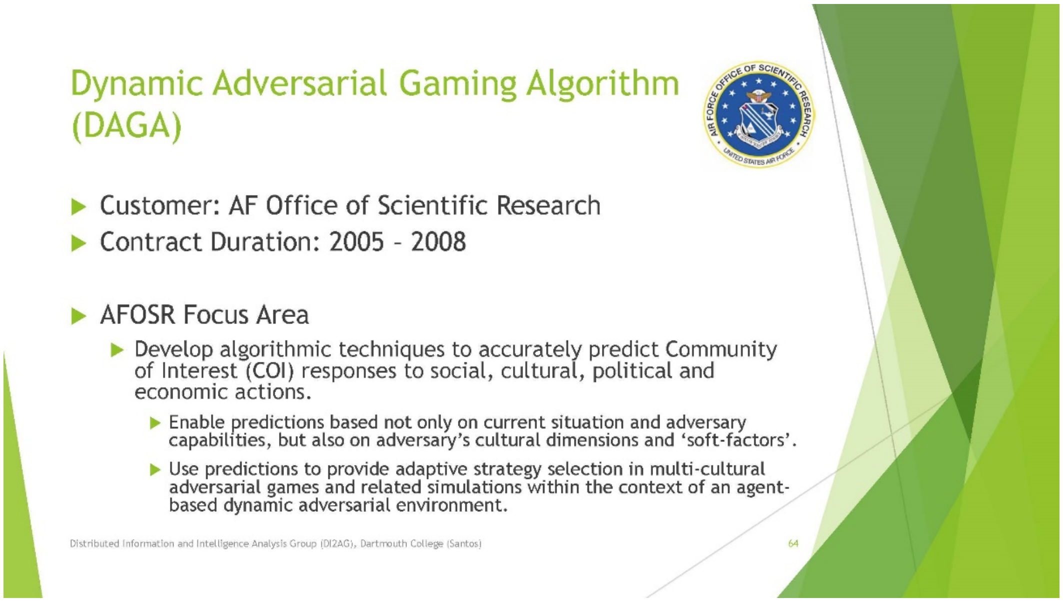
Simulation Results

- Aggressive commander had higher likelihood to actively respond
- Passive commander had higher likelihood to merely defend

- Passive commander caused more damage to blue forces
  - Preserved assets at beginning by shutting down all equipment (SAMs, radars, etc.)
    - Harder targeting problem for Blue forces bombers
  - Red commander's decision to attack occurred later when Blue aircraft were turning around and disengaging
    - Blue aircraft were in disadvantageous firing position to handle Red SAMs, etc.

- Conclusions
  - Rapid assembly of adversary intent models through domain ontology
  - Parametric-driven change in intent based on soft factors such as aggressiveness
  - Dynamic behavior change/response based on sequence of red-blue interactions such as depletion of resources

Lehman, Lynn A., Krause, Lee S., Gilmour, Duane A., Santos, Eugene, Jr., and Zhao, Qunhua, "Intent Driven Adversarial Modeling," *Proceedings of the Tenth International Command and Control Research and Technology Symposium: The Future of C2*, McLean, VA, 2005

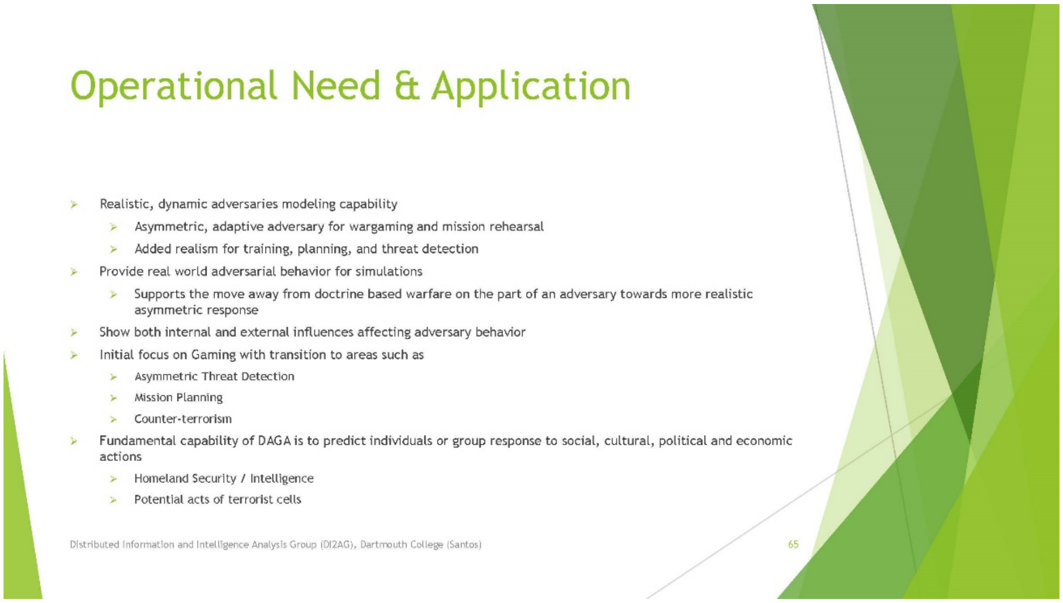


Dynamic Adversarial Gaming Algorithm (DAGA)

- Customer: AF Office of Scientific Research
- Contract Duration: 2005 – 2008

- AFOSR Focus Area
  - Develop algorithmic techniques to accurately predict Community of Interest (COI) responses to social, cultural, political and economic actions.
    - Enable predictions based not only on current situation and adversary capabilities, but also on adversary's cultural dimensions and 'soft-factors'.
    - Use predictions to provide adaptive strategy selection in multi-cultural adversarial games and related simulations within the context of an agent-based dynamic adversarial environment.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 64



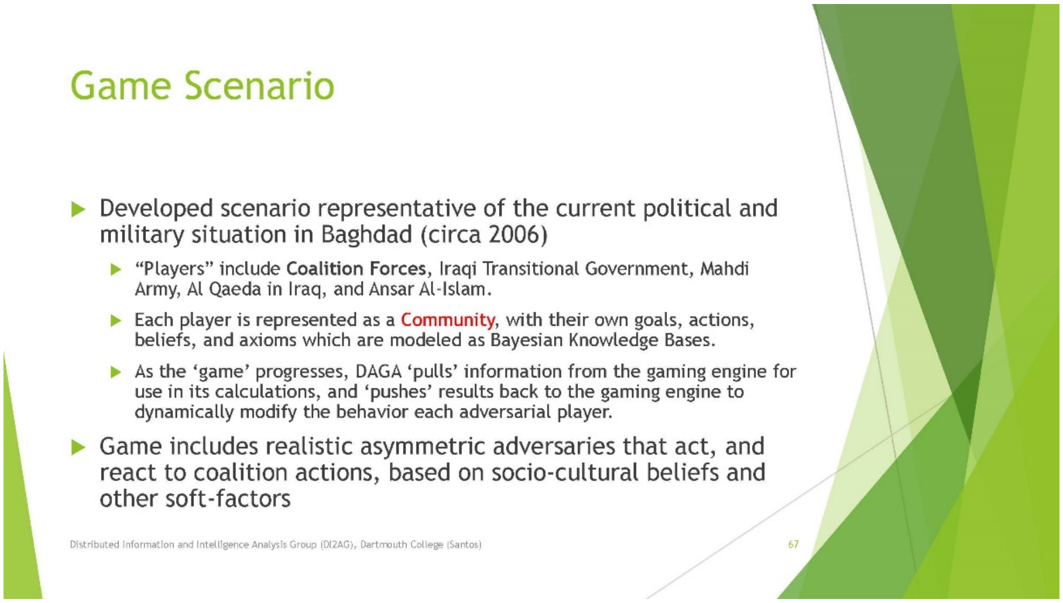
Operational Need & Application

- Realistic, dynamic adversaries modeling capability
  - Asymmetric, adaptive adversary for wargaming and mission rehearsal
  - Added realism for training, planning, and threat detection
- Provide real world adversarial behavior for simulations
  - Supports the move away from doctrine based warfare on the part of an adversary towards more realistic asymmetric response
- Show both internal and external influences affecting adversary behavior
- Initial focus on Gaming with transition to areas such as
  - Asymmetric Threat Detection
  - Mission Planning
  - Counter-terrorism
- Fundamental capability of DAGA is to predict individuals or group response to social, cultural, political and economic actions
  - Homeland Security / Intelligence
  - Potential acts of terrorist cells

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 65

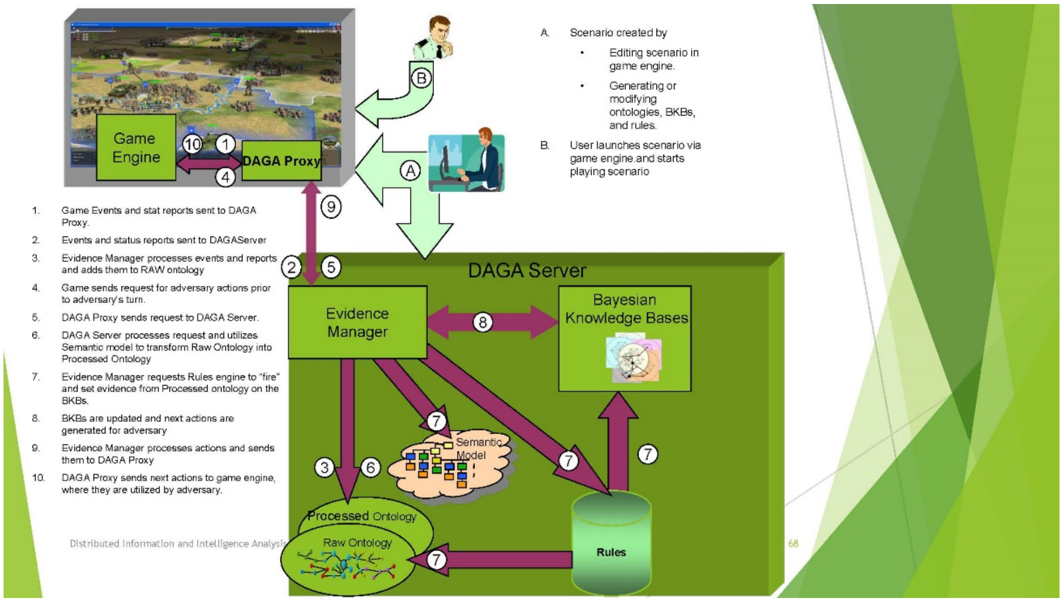
Game Integration

- To highlight DAGA's capabilities, we have integrated it with the popular *Civilization 4* (Civ4) game engine to demonstrate how the infusion of socio-cultural influences leads to a much more realistic asymmetric adversary.



Game Scenario

- Developed scenario representative of the current political and military situation in Baghdad (circa 2006)
  - "Players" include **Coalition Forces**, Iraqi Transitional Government, Mahdi Army, Al Qaeda in Iraq, and Ansar Al-Islam.
  - Each player is represented as a Community, with their own goals, actions, beliefs, and axioms which are modeled as Bayesian Knowledge Bases.
  - As the 'game' progresses, DAGA 'pulls' information from the gaming engine for use in its calculations, and 'pushes' results back to the gaming engine to dynamically modify the behavior each adversarial player.
- Game includes realistic asymmetric adversaries that act, and react to coalition actions, based on socio-cultural beliefs and other soft-factors



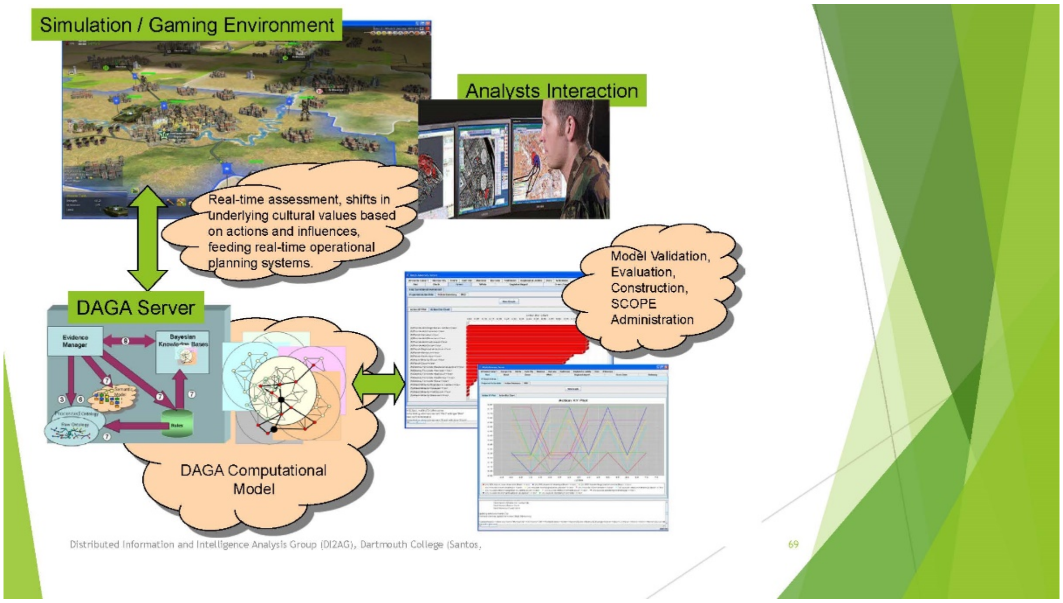
A. Scenario created by
- Editing scenario in game engine.
- Generating or modifying ontologies, BKBs, and rules.

B. User launches scenario via game engine and starts playing scenario

1. Game Events and stat reports sent to DAGA Proxy.
2. Events and status reports sent to DAGAServer
3. Evidence Manager processes events and reports and adds them to RAW ontology
4. Game sends request for adversary actions prior to adversary's turn.
5. DAGA Proxy sends request to DAGA Server.
6. DAGA Server processes request and utilizes Semantic model to transform Raw Ontology into Processed Ontology
7. Evidence Manager requests Rules engine to "fire" and set evidence from Processed ontology on the BKBs.
8. BKBs are updated and next actions are generated for adversary
9. Evidence Manager processes actions and sends them to DAGA Proxy
10. DAGA Proxy sends next actions to game engine, where they are utilized by adversary.

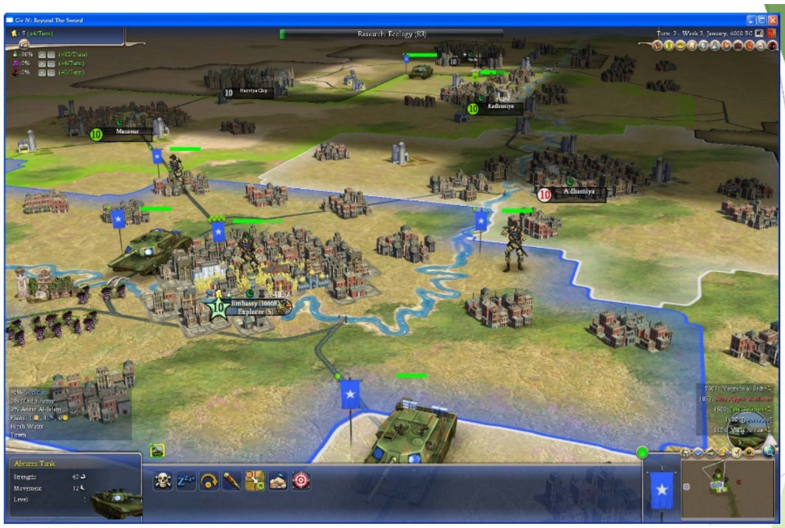
Distributed Information and Intelligence Analysis

68


Simulation / Gaming Environment
Analysts Interaction
Real-time assessment, shifts in underlying cultural values based on actions and influences, feeding real-time operational planning systems.
DAGA Server
Model Validation, Evaluation, Construction, SCOPE Administration
DAGA Computational Model

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos,

69

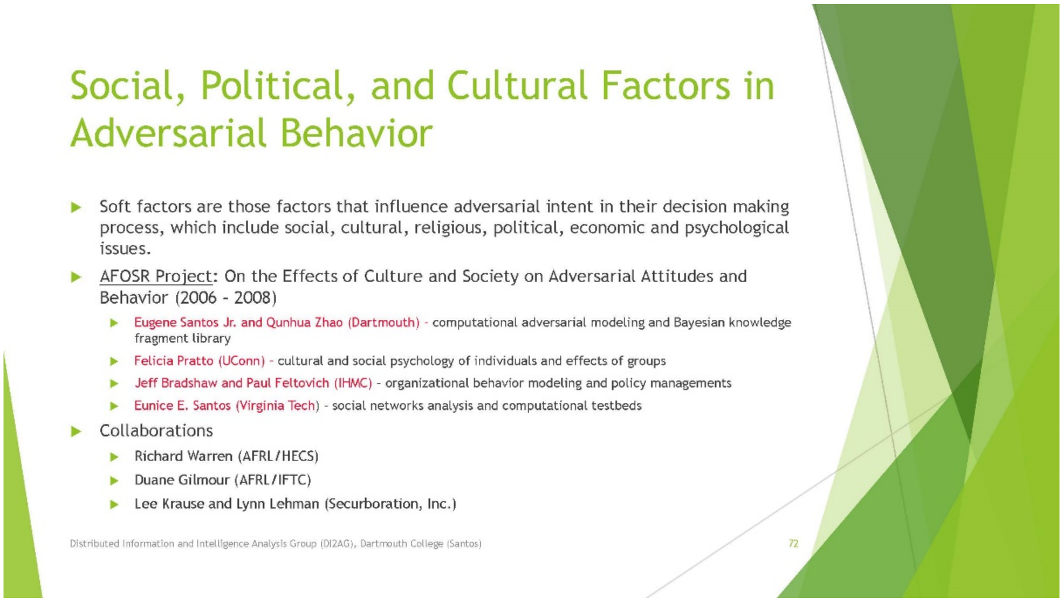


Groups of adversaries and neutrals being driven by DAGA – reacting to coalition actions based on the current game state and their goals, internal beliefs, external beliefs, and actions.

Santos, Eugene, Jr., Zhao, Qunhua, Pratto, Felicia, Pearson, Adam R., McQueary, Bruce, Breeden, Andy, and Krause, Lee, "Modeling Multiple Communities of Interest for Interactive Simulation and Gaming: The Dynamic Adversarial Gaming Algorithm Project," *Proceedings of the SPIE: Defense & Security Symposium*, Vol. 6564, Orlando, FL, 2007.

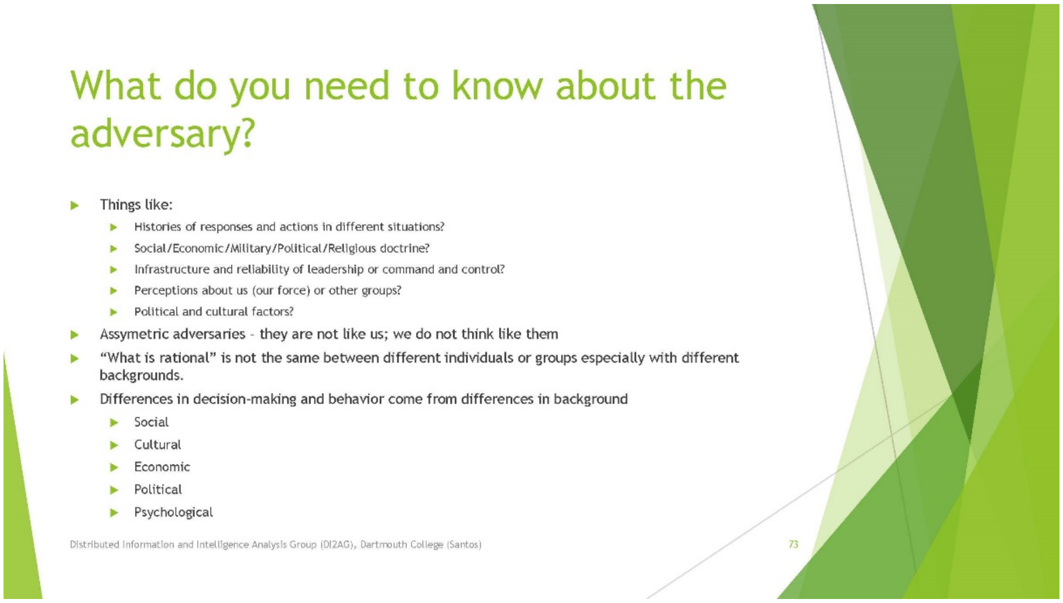
Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models



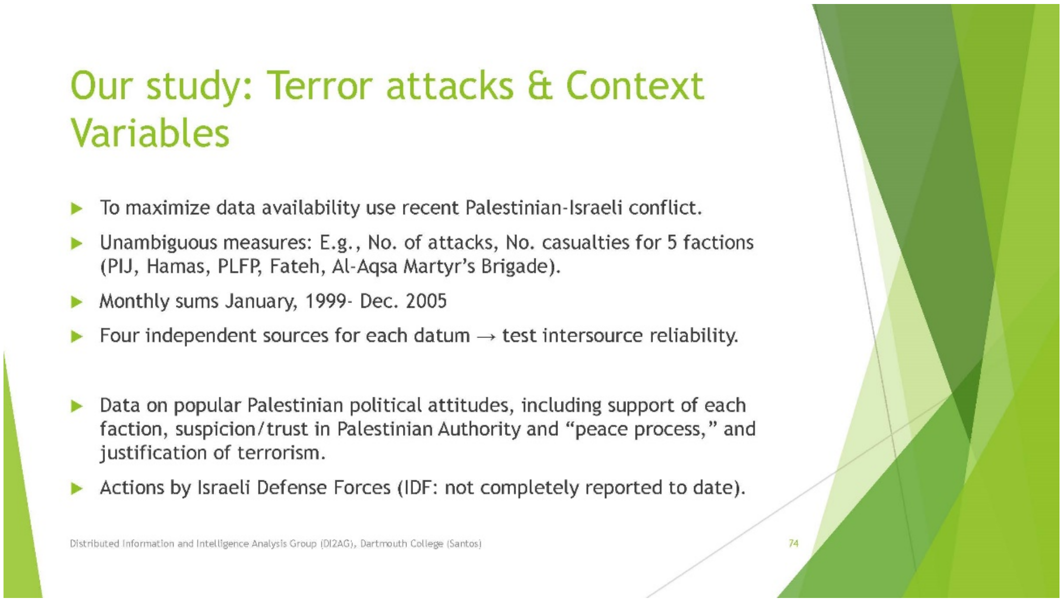
Social, Political, and Cultural Factors in Adversarial Behavior

- Soft factors are those factors that influence adversarial intent in their decision making process, which include social, cultural, religious, political, economic and psychological issues.
- AFOSR Project: On the Effects of Culture and Society on Adversarial Attitudes and Behavior (2006 - 2008)
  - Eugene Santos Jr. and Qunhua Zhao (Dartmouth) - computational adversarial modeling and Bayesian knowledge fragment library
  - Felicia Pratto (UConn) - cultural and social psychology of individuals and effects of groups
  - Jeff Bradshaw and Paul Feltovich (IHMC) - organizational behavior modeling and policy managements
  - Eunice E. Santos (Virginia Tech) - social networks analysis and computational testbeds
- Collaborations
  - Richard Warren (AFRL/HECS)
  - Duane Gilmour (AFRL/IFTC)
  - Lee Krause and Lynn Lehman (Securboration, Inc.)



What do you need to know about the adversary?

- Things like:
  - Histories of responses and actions in different situations?
  - Social/Economic/Military/Political/Religious doctrine?
  - Infrastructure and reliability of leadership or command and control?
  - Perceptions about us (our force) or other groups?
  - Political and cultural factors?
- Assymetric adversaries - they are not like us; we do not think like them
- "What is rational" is not the same between different individuals or groups especially with different backgrounds.
- Differences in decision-making and behavior come from differences in background
  - Social
  - Cultural
  - Economic
  - Political
  - Psychological

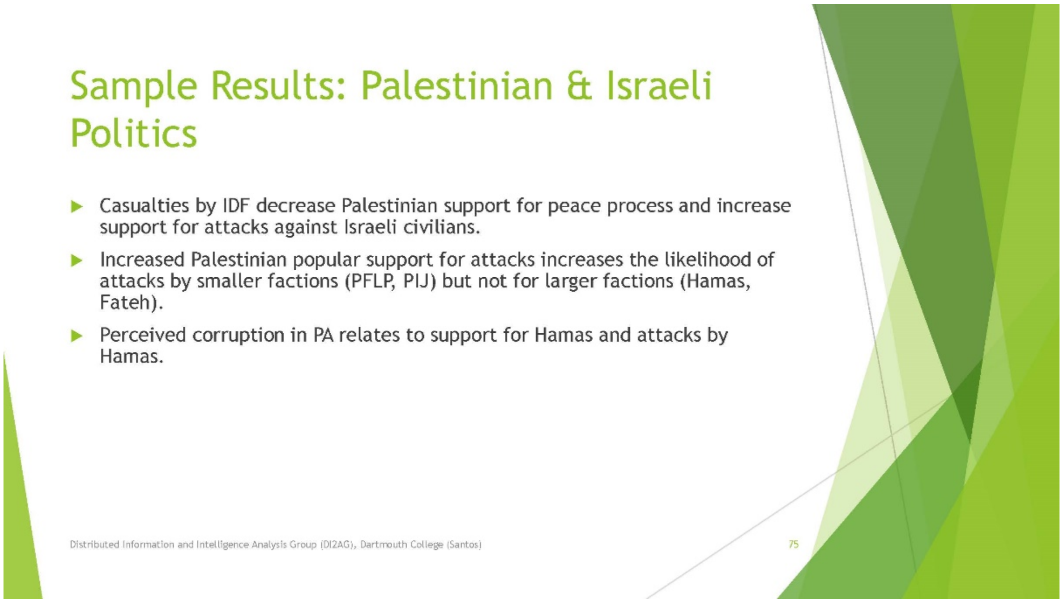


Our study: Terror attacks & Context Variables

- To maximize data availability use recent Palestinian-Israeli conflict.
- Unambiguous measures: E.g., No. of attacks, No. casualties for 5 factions (PIJ, Hamas, PLFP, Fateh, Al-Aqsa Martyr's Brigade).
- Monthly sums January, 1999- Dec. 2005
- Four independent sources for each datum → test intersource reliability.

- Data on popular Palestinian political attitudes, including support of each faction, suspicion/trust in Palestinian Authority and "peace process," and justification of terrorism.
- Actions by Israeli Defense Forces (IDF: not completely reported to date).

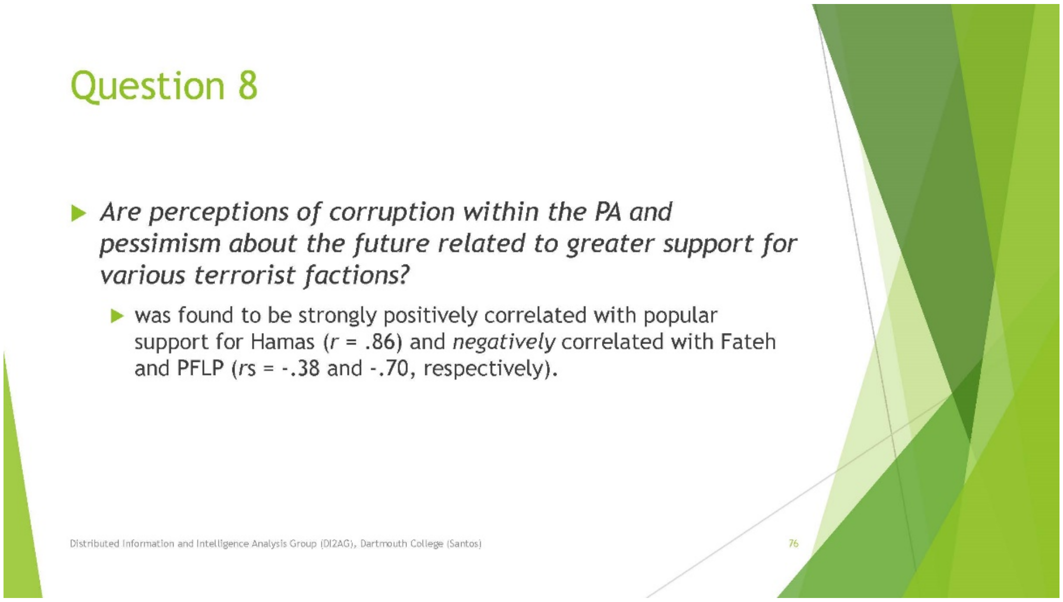
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 74



Sample Results: Palestinian & Israeli Politics

- Casualties by IDF decrease Palestinian support for peace process and increase support for attacks against Israeli civilians.
- Increased Palestinian popular support for attacks increases the likelihood of attacks by smaller factions (PFLP, PIJ) but not for larger factions (Hamas, Fateh).
- Perceived corruption in PA relates to support for Hamas and attacks by Hamas.

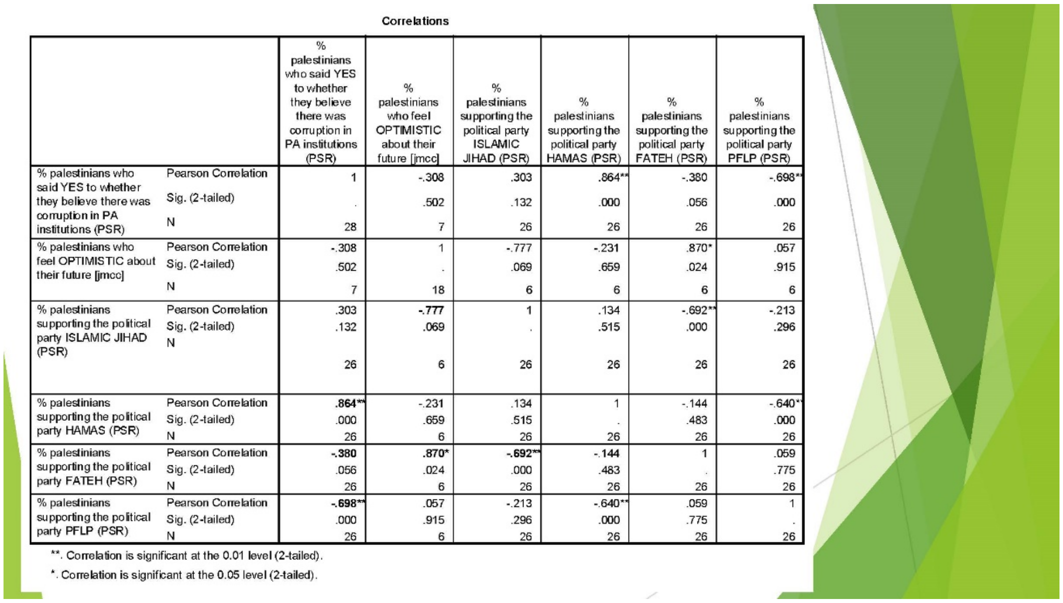
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 75



Question 8

- *Are perceptions of corruption within the PA and pessimism about the future related to greater support for various terrorist factions?*
    - was found to be strongly positively correlated with popular support for Hamas (*r* = .86) and *negatively* correlated with Fateh and PFLP (*r*s = -.38 and -.70, respectively).

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 76

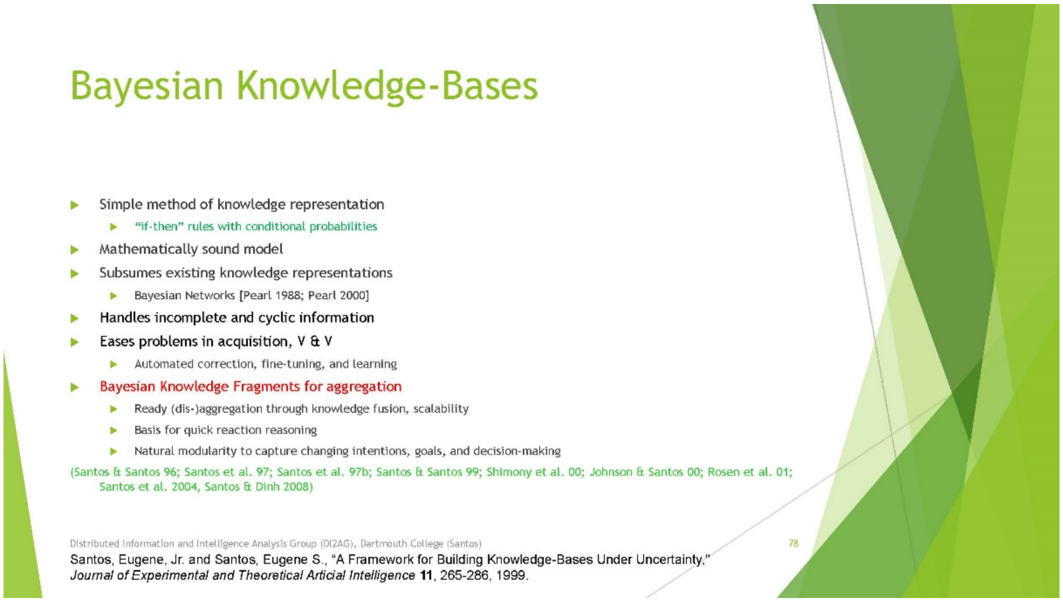


Correlations

		% palestinians who said YES to whether they believe there was corruption in PA institutions (PSR)	% palestinians who feel OPTIMISTIC about their future [jmcc]	% palestinians supporting the political party ISLAMIC JIHAD (PSR)	% palestinians supporting the political party HAMAS (PSR)	% palestinians supporting the political party FATEH (PSR)	% palestinians supporting the political party PFLP (PSR)
% palestinians who said YES to whether they believe there was corruption in PA institutions (PSR)	Pearson Correlation	1	-.308	.303	.864**	-.380	-.698**
	Sig. (2-tailed)	.	.502	.132	.000	.056	.000
	N	28	7	26	26	26	26
% palestinians who feel OPTIMISTIC about their future [jmcc]	Pearson Correlation	-.308	1	-.777	-.231	.870*	.057
	Sig. (2-tailed)	.502	.	.069	.659	.024	.915
	N	7	18	6	6	6	6
% palestinians supporting the political party ISLAMIC JIHAD (PSR)	Pearson Correlation	.303	**-.777**	1	.134	-.692**	-.213
	Sig. (2-tailed)	.132	.069	.	.515	.000	.296
	N	26	6	26	26	26	26
% palestinians supporting the political party HAMAS (PSR)	Pearson Correlation	**.864****	-.231	.134	1	-.144	-.640**
	Sig. (2-tailed)	.000	.659	.515	.	.483	.000
	N	26	6	26	26	26	26
% palestinians supporting the political party FATEH (PSR)	Pearson Correlation	**-.380**	**.870***	**-.692****	-.144	1	.059
	Sig. (2-tailed)	.056	.024	.000	.483	.	.775
	N	26	6	26	26	26	26
% palestinians supporting the political party PFLP (PSR)	Pearson Correlation	**-.698****	.057	-.213	-.640**	.059	1
	Sig. (2-tailed)	.000	.915	.296	.000	.775	.
	N	26	6	26	26	26	26

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).



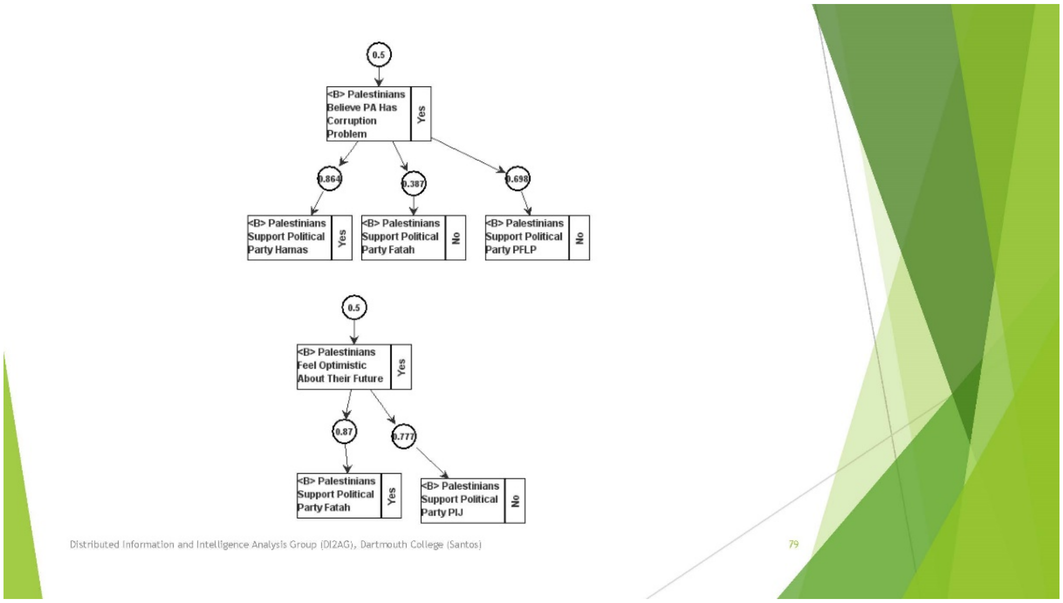
Bayesian Knowledge-Bases

- Simple method of knowledge representation
  - "if-then" rules with conditional probabilities
- Mathematically sound model
- Subsumes existing knowledge representations
  - Bayesian Networks [Pearl 1988; Pearl 2000]
- Handles incomplete and cyclic information
- Eases problems in acquisition, V & V
  - Automated correction, fine-tuning, and learning
- Bayesian Knowledge Fragments for aggregation
  - Ready (dis-)aggregation through knowledge fusion, scalability
  - Basis for quick reaction reasoning
  - Natural modularity to capture changing intentions, goals, and decision-making

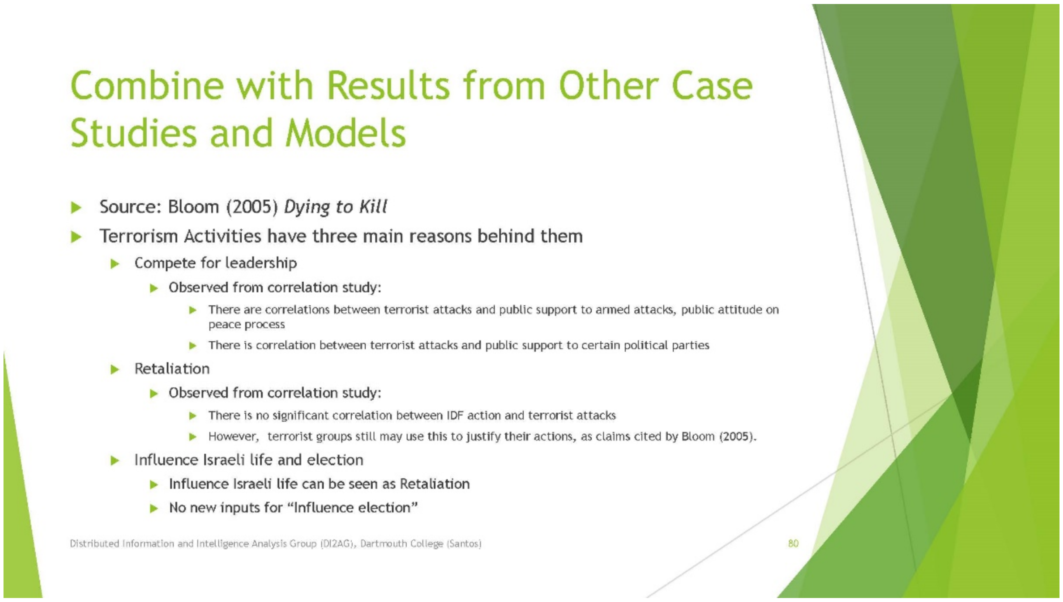
(Santos & Santos 96; Santos et al. 97; Santos et al. 97b; Santos & Santos 99; Shimony et al. 00; Johnson & Santos 00; Rosen et al. 01; Santos et al. 2004, Santos & Dinh 2008)

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 78

Santos, Eugene, Jr. and Santos, Eugene S., "A Framework for Building Knowledge-Bases Under Uncertainty," *Journal of Experimental and Theoretical Articial Intelligence* **11**, 265-286, 1999.


0.5
<B> Palestinians Believe PA Has Corruption Problem
Yes
0.864
0.387
0.698
<B> Palestinians Support Political Party Hamas
Yes
<B> Palestinians Support Political Party Fatah
No
<B> Palestinians Support Political Party PFLP
No
0.5
<B> Palestinians Feel Optimistic About Their Future
Yes
0.87
0.777
<B> Palestinians Support Political Party Fatah
Yes
<B> Palestinians Support Political Party PIJ
No

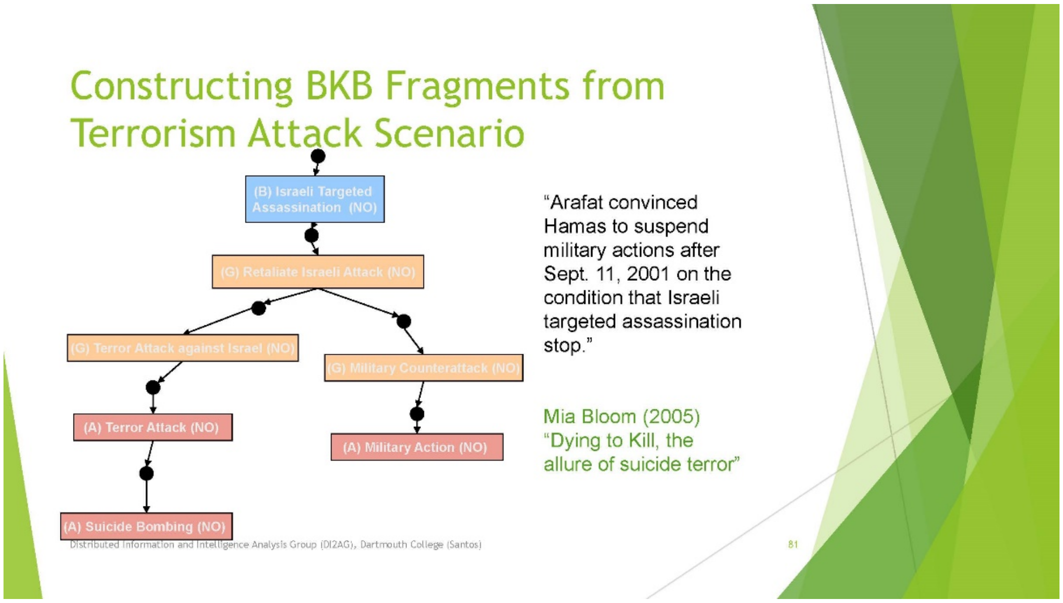
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 79



Combine with Results from Other Case Studies and Models

- Source: Bloom (2005) *Dying to Kill*
- Terrorism Activities have three main reasons behind them
  - Compete for leadership
    - Observed from correlation study:
      - There are correlations between terrorist attacks and public support to armed attacks, public attitude on peace process
      - There is correlation between terrorist attacks and public support to certain political parties
  - Retaliation
    - Observed from correlation study:
      - There is no significant correlation between IDF action and terrorist attacks
      - However, terrorist groups still may use this to justify their actions, as claims cited by Bloom (2005).
  - Influence Israeli life and election
    - Influence Israeli life can be seen as Retaliation
    - No new inputs for "Influence election"

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 80



Constructing BKB Fragments from Terrorism Attack Scenario

(B) Israeli Targeted Assassination (NO)

(G) Retaliate Israeli Attack (NO)

(G) Terror Attack against Israel (NO)

(G) Military Counterattack (NO)

(A) Terror Attack (NO)

(A) Military Action (NO)

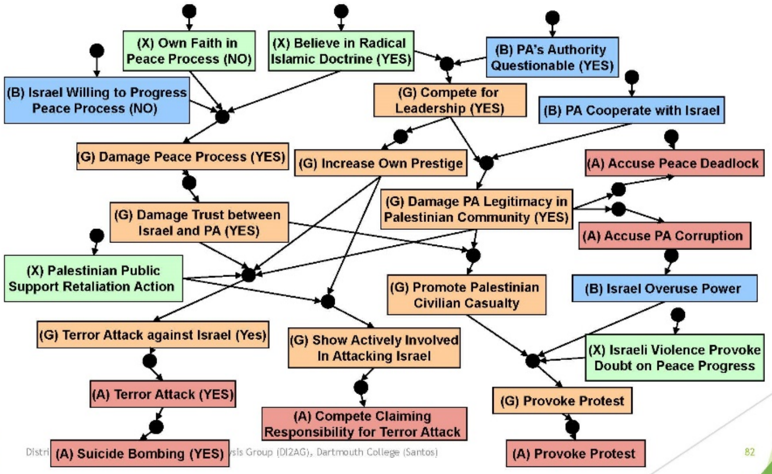
(A) Suicide Bombing (NO)

"Arafat convinced Hamas to suspend military actions after Sept. 11, 2001 on the condition that Israeli targeted assassination stop."

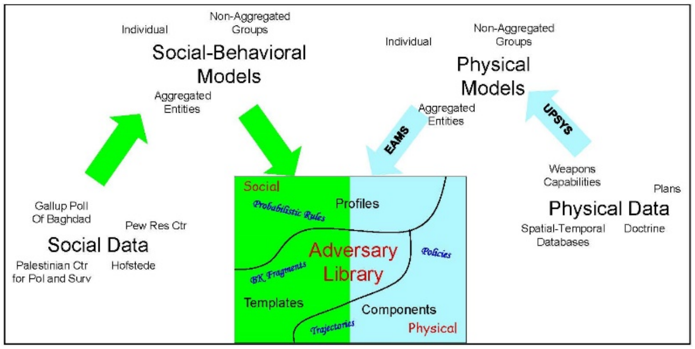
Mia Bloom (2005) "Dying to Kill, the allure of suicide terror"

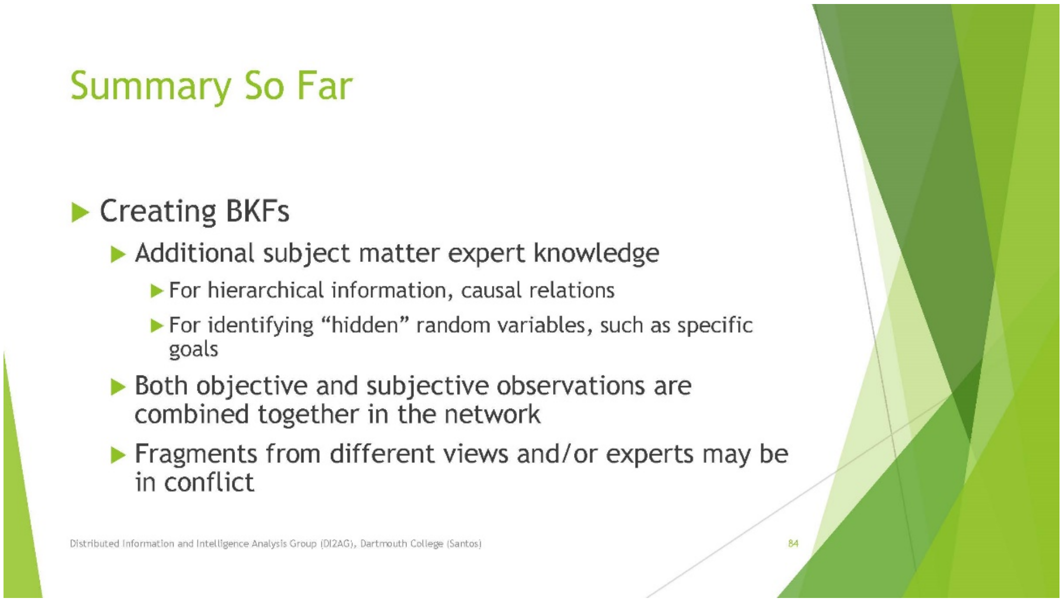
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 81

Another view of the reason behind suicide bombing: Competing for the leadership in Palestinian community, when public has no hope in peace and supports violence for revenge.
(1) Increasing own profile; (2) damage PA's authority; and (3) damage peace process


(X) Own Faith in Peace Process (NO)
(X) Believe in Radical Islamic Doctrine (YES)
(B) PA's Authority Questionable (YES)
(B) Israel Willing to Progress Peace Process (NO)
(G) Compete for Leadership (YES)
(B) PA Cooperate with Israel
(G) Damage Peace Process (YES)
(G) Increase Own Prestige
(A) Accuse Peace Deadlock
(G) Damage Trust between Israel and PA (YES)
(G) Damage PA Legitimacy in Palestinian Community (YES)
(A) Accuse PA Corruption
(X) Palestinian Public Support Retaliation Action
(G) Promote Palestinian Civilian Casualty
(B) Israel Overuse Power
(G) Terror Attack against Israel (Yes)
(G) Show Actively Involved In Attacking Israel
(X) Israeli Violence Provoke Doubt on Peace Progress
(A) Terror Attack (YES)
(A) Compete Claiming Responsibility for Terror Attack
(G) Provoke Protest
(A) Suicide Bombing (YES)
(A) Provoke Protest

Adversary Fragment Library

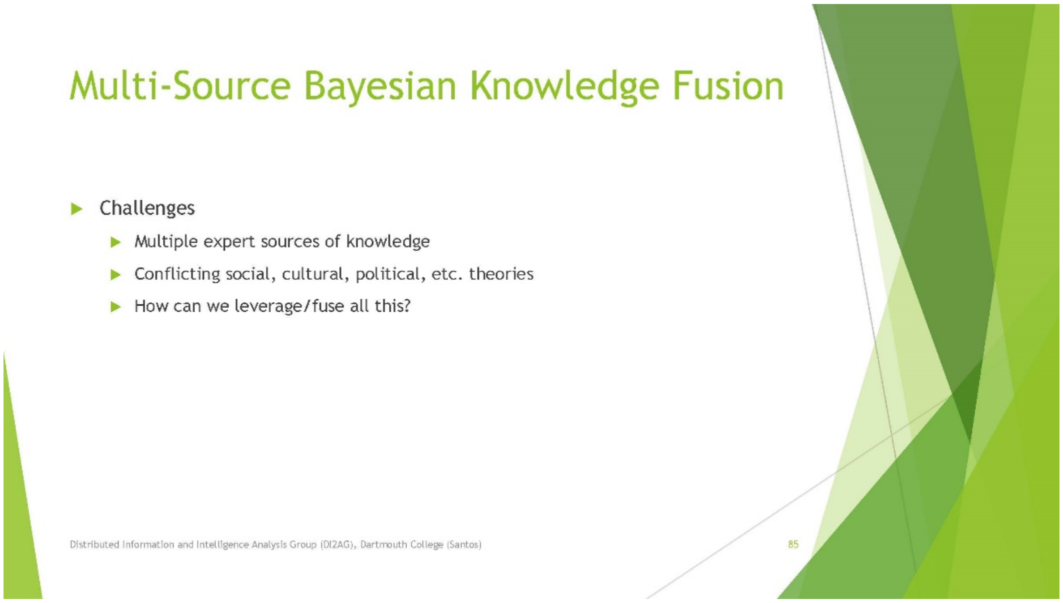




Summary So Far

- Creating BKFs
  - Additional subject matter expert knowledge
    - For hierarchical information, causal relations
    - For identifying “hidden” random variables, such as specific goals
  - Both objective and subjective observations are combined together in the network
  - Fragments from different views and/or experts may be in conflict

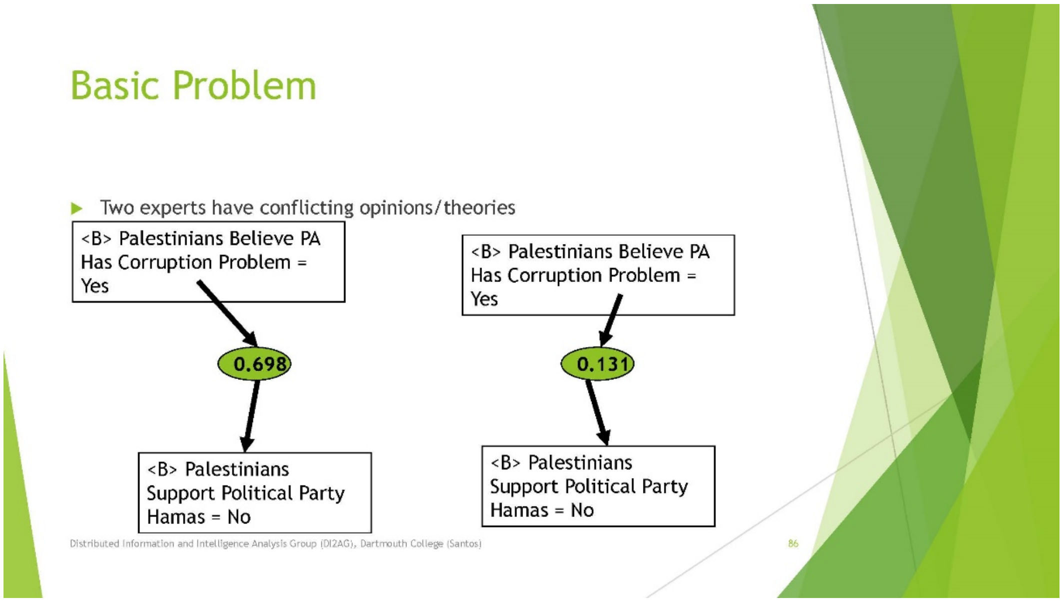
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 84



Multi-Source Bayesian Knowledge Fusion

- Challenges
  - Multiple expert sources of knowledge
  - Conflicting social, cultural, political, etc. theories
  - How can we leverage/fuse all this?

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 85



Basic Problem

- Two experts have conflicting opinions/theories

<B> Palestinians Believe PA Has Corruption Problem = Yes

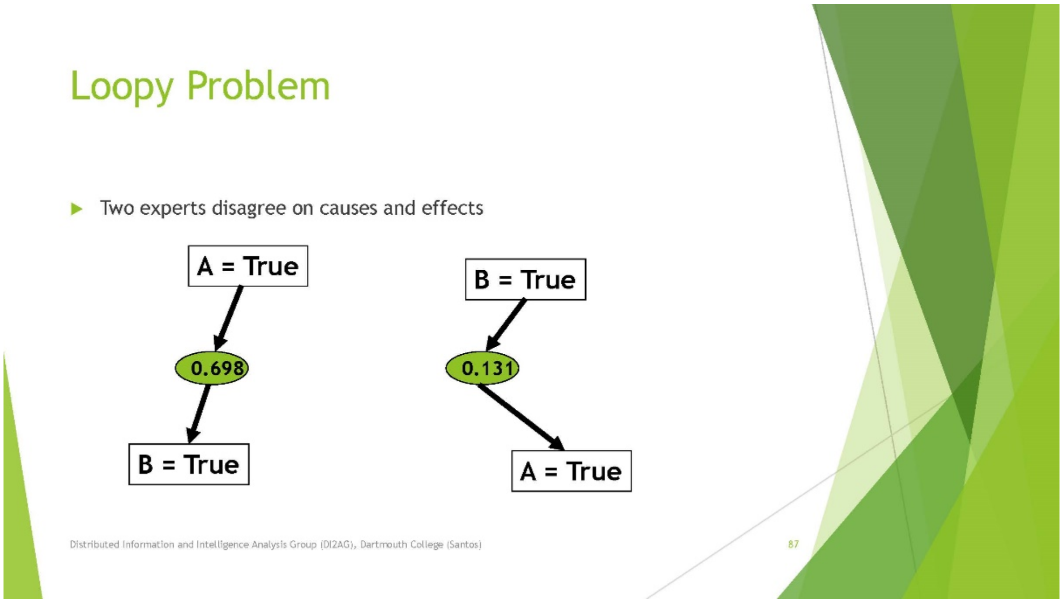
0.698

<B> Palestinians Support Political Party Hamas = No

<B> Palestinians Believe PA Has Corruption Problem = Yes

0.131

<B> Palestinians Support Political Party Hamas = No



Loopy Problem

- Two experts disagree on causes and effects

A = True

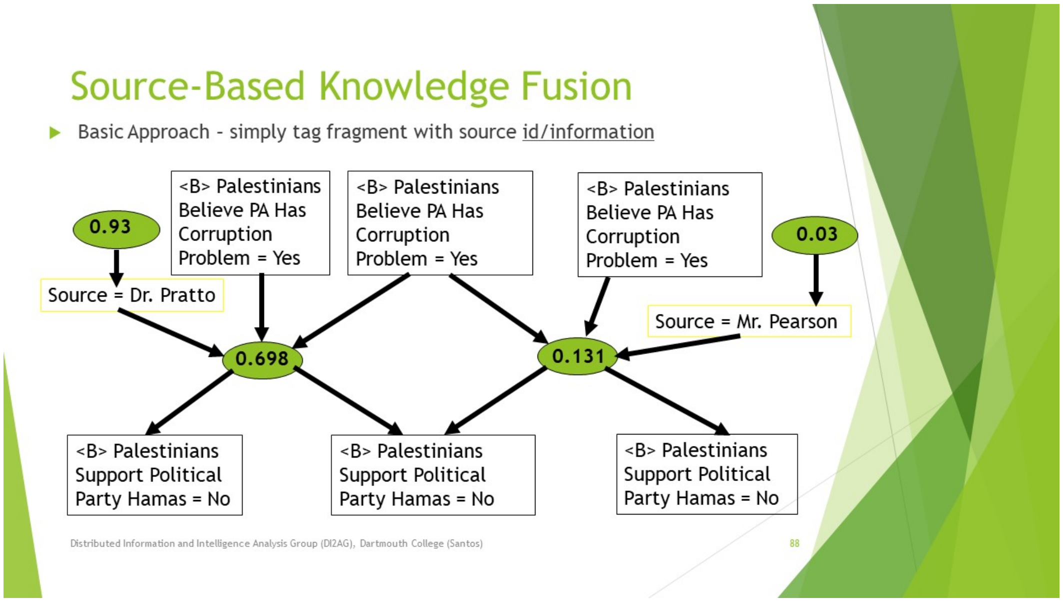
0.698

B = True

B = True

0.131

A = True



Source-Based Knowledge Fusion

- Basic Approach - simply tag fragment with source id/information

0.93

Source = Dr. Pratto

<B> Palestinians Believe PA Has Corruption Problem = Yes

<B> Palestinians Believe PA Has Corruption Problem = Yes

<B> Palestinians Believe PA Has Corruption Problem = Yes

0.03

Source = Mr. Pearson

0.698

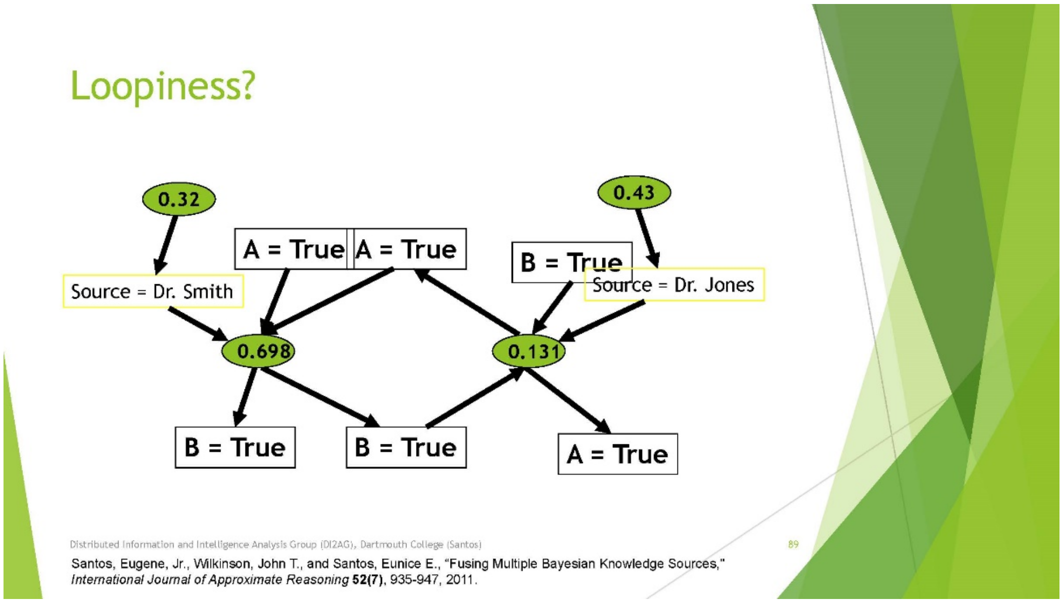
0.131

<B> Palestinians Support Political Party Hamas = No

<B> Palestinians Support Political Party Hamas = No

<B> Palestinians Support Political Party Hamas = No

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 88



Loopiness?

0.32

Source = Dr. Smith

A = True

A = True

B = True

0.43

Source = Dr. Jones

0.698

0.131

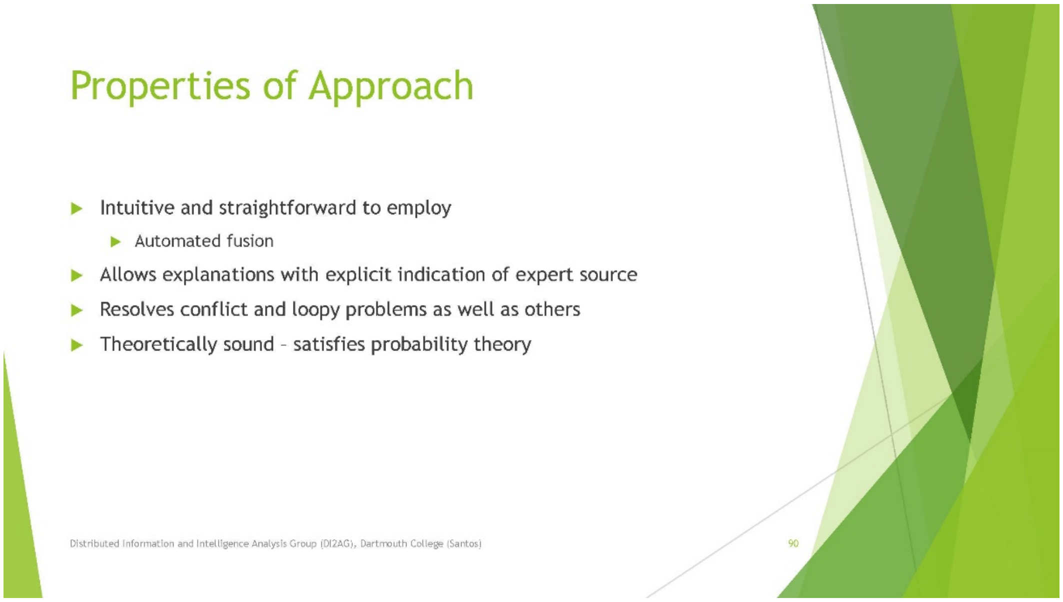
B = True

B = True

A = True

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 89

Santos, Eugene, Jr., Wilkinson, John T., and Santos, Eunice E., "Fusing Multiple Bayesian Knowledge Sources," *International Journal of Approximate Reasoning* **52(7)**, 935-947, 2011.

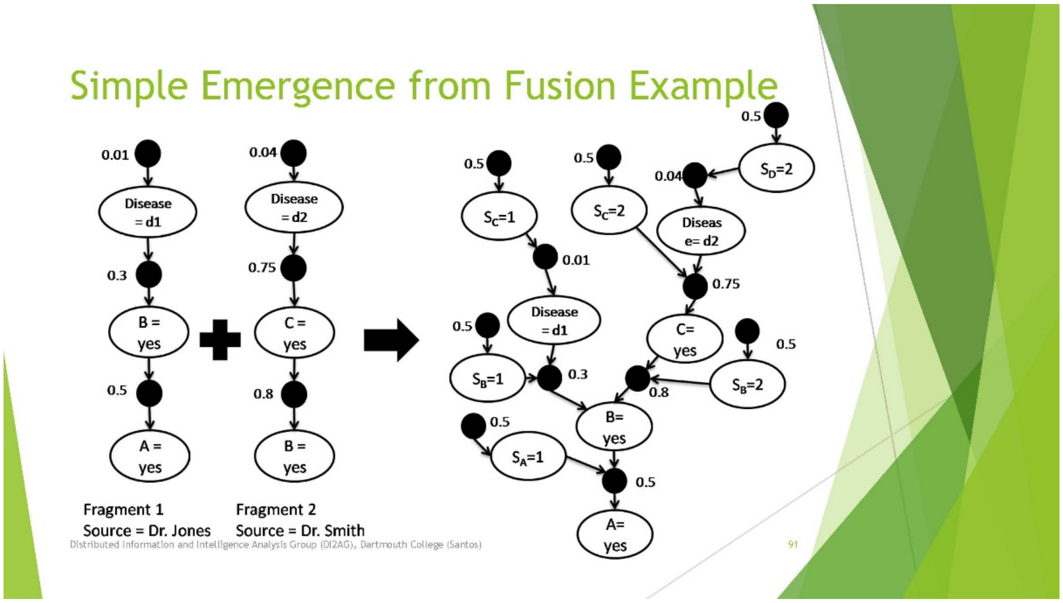


Properties of Approach

- Intuitive and straightforward to employ
  - Automated fusion
- Allows explanations with explicit indication of expert source
- Resolves conflict and loopy problems as well as others
- Theoretically sound – satisfies probability theory

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

90



Simple Emergence from Fusion Example

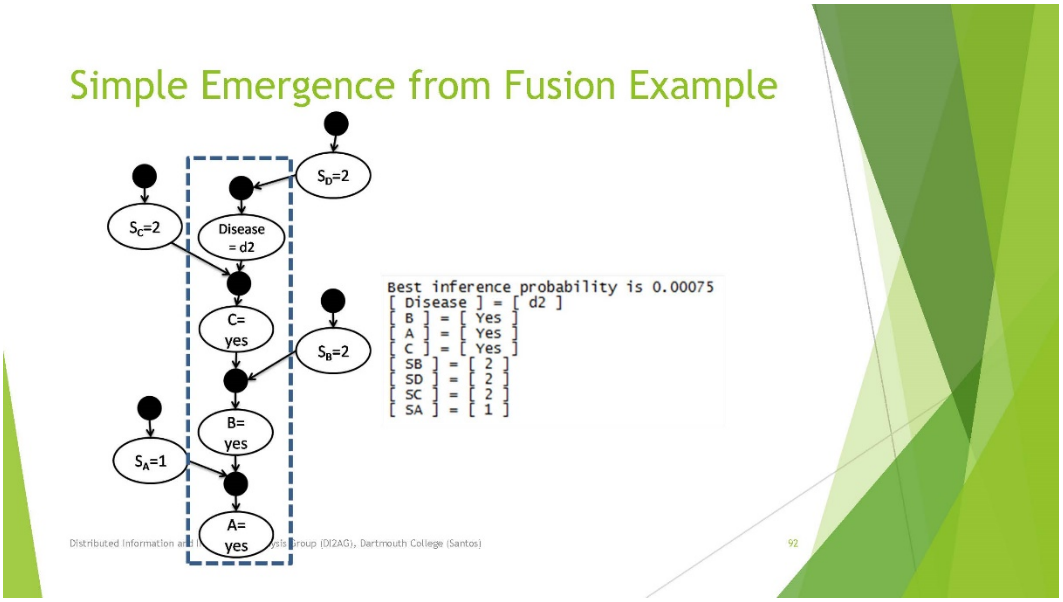
Fragment 1
Source = Dr. Jones

Fragment 2
Source = Dr. Smith

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

91

Simple Emergence from Fusion Example

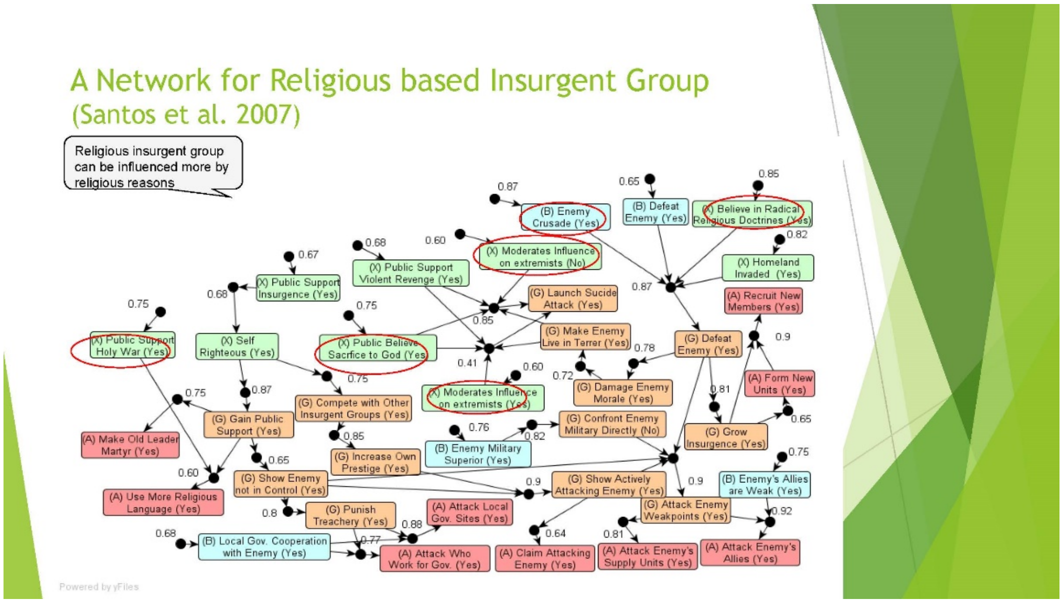


```
Best inference probability is 0.00075
[ Disease ] = [ d2 ]
[ B ] = [ Yes ]
[ A ] = [ Yes ]
[ C ] = [ Yes ]
[ SB ] = [ 2 ]
[ SD ] = [ 2 ]
[ SC ] = [ 2 ]
[ SA ] = [ 1 ]
```

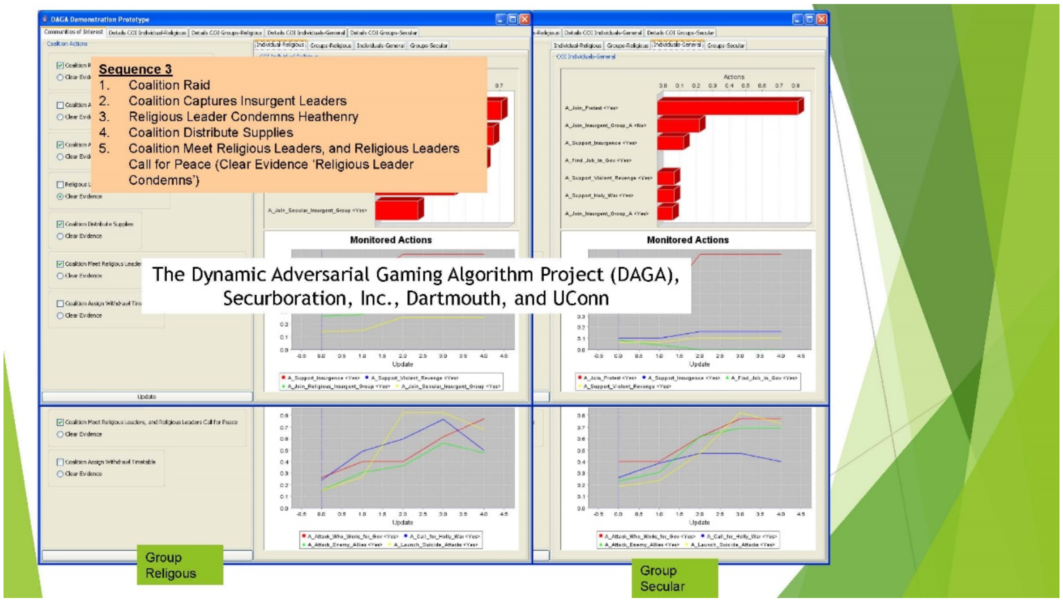
Distributed Information and I ysis Group (DI2AG), Dartmouth College (Santos)

92

A Network for Religious based Insurgent Group (Santos et al. 2007)


Religious insurgent group can be influenced more by religious reasons

0.87
(B) Enemy Crusade (Yes)
0.65
(B) Defeat Enemy (Yes)
0.85
(X) Believe in Radical Religious Doctrines (Yes)
0.82
(X) Homeland Invaded (Yes)
0.68
0.60
(X) Moderates Influence on extremists (No)
0.67
(X) Public Support Violent Revenge (Yes)
(X) Public Support Insurgence (Yes)
0.68
0.87
(A) Recruit New Members (Yes)
(G) Launch Sucide Attack (Yes)
0.75
0.75
0.85
(G) Make Enemy Live in Terrer (Yes)
0.9
(X) Public Support Holy War (Yes)
(X) Self Righteous (Yes)
(X) Public Believe Sacrfice to God (Yes)
0.78
(G) Defeat Enemy (Yes)
0.41
0.60
0.72
0.75
(A) Form New Units (Yes)
0.75
0.87
(X) Moderates Influence on extremists (Yes)
(G) Damage Enemy Morale (Yes)
0.81
(G) Compete with Other Insurgent Groups (Yes)
(G) Gain Public Support (Yes)
(G) Confront Enemy Military Directly (No)
0.65
(A) Make Old Leader Martyr (Yes)
0.76
0.82
(G) Grow Insurgence (Yes)
0.85
(B) Enemy Military Superior (Yes)
0.65
(G) Increase Own Prestige (Yes)
0.75
0.60
(G) Show Enemy not in Control (Yes)
0.9
(G) Show Actively Attacking Enemy (Yes)
0.9
(B) Enemy's Allies are Weak (Yes)
(A) Use More Religious Language (Yes)
(G) Attack Enemy Weakpoints (Yes)
0.8
(G) Punish Treachery (Yes)
(A) Attack Local Gov. Sites (Yes)
0.92
0.88
0.68
(B) Local Gov. Cooperation with Enemy (Yes)
0.77
0.64
0.81
(A) Attack Who Work for Gov. (Yes)
(A) Claim Attacking Enemy (Yes)
(A) Attack Enemy's Supply Units (Yes)
(A) Attack Enemy's Allies (Yes)
Powered by yFiles



Sequence 3

1. Coalition Raid
2. Coalition Captures Insurgent Leaders
3. Religious Leader Condemns Heathenry
4. Coalition Distribute Supplies
5. Coalition Meet Religious Leaders, and Religious Leaders Call for Peace (Clear Evidence 'Religious Leader Condemns')

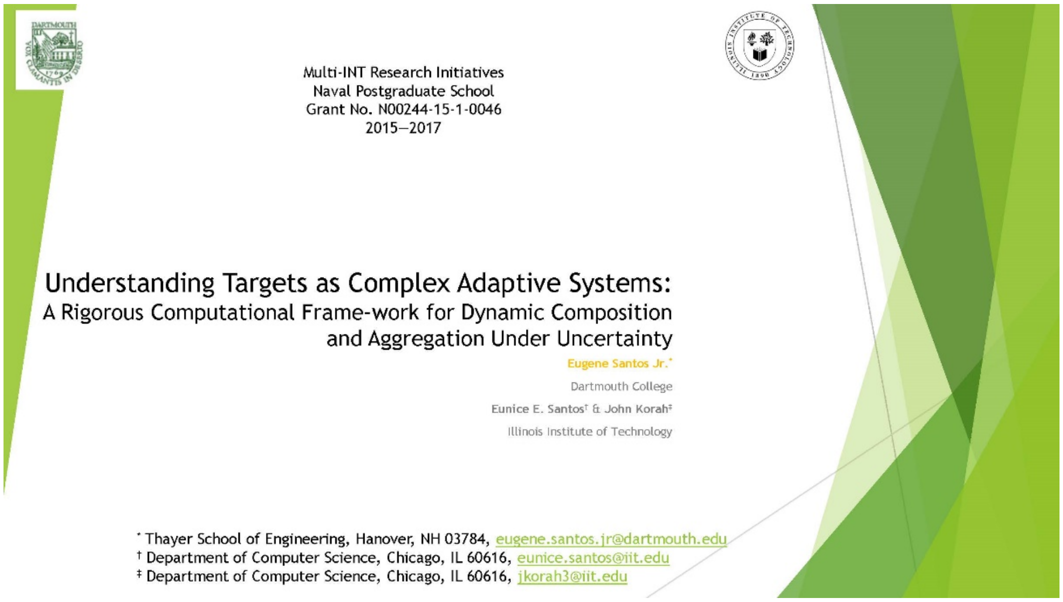
The Dynamic Adversarial Gaming Algorithm Project (DAGA), Securboration, Inc., Dartmouth, and UConn

Group Religous

Group Secular

Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models



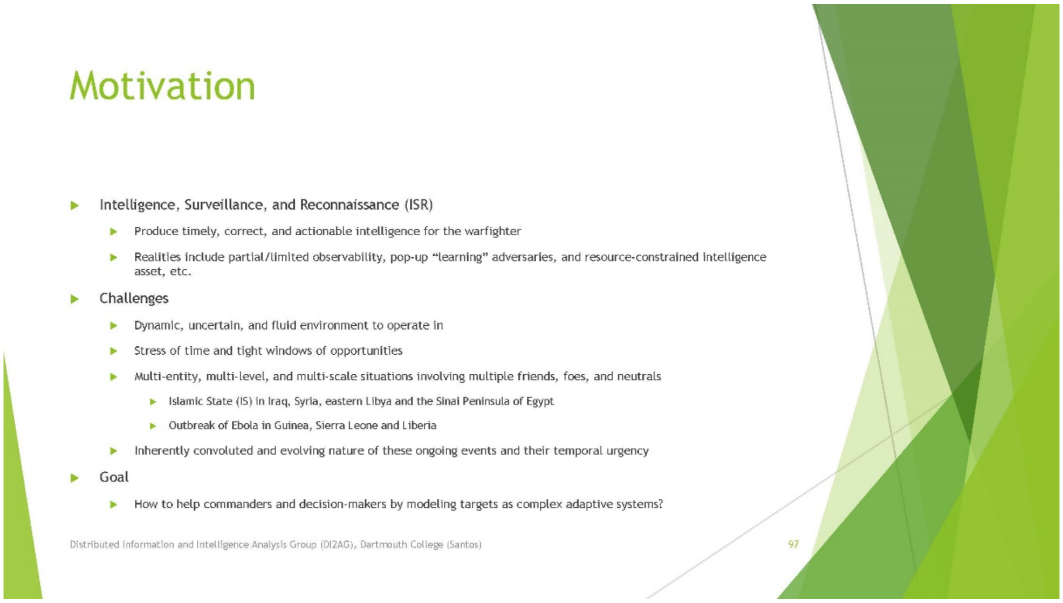
Multi-INT Research Initiatives
Naval Postgraduate School
Grant No. N00244-15-1-0046
2015—2017

Understanding Targets as Complex Adaptive Systems:

A Rigorous Computational Frame-work for Dynamic Composition and Aggregation Under Uncertainty

Eugene Santos Jr.[*]
Dartmouth College
Eunice E. Santos[†] & John Korah[‡]
Illinois Institute of Technology

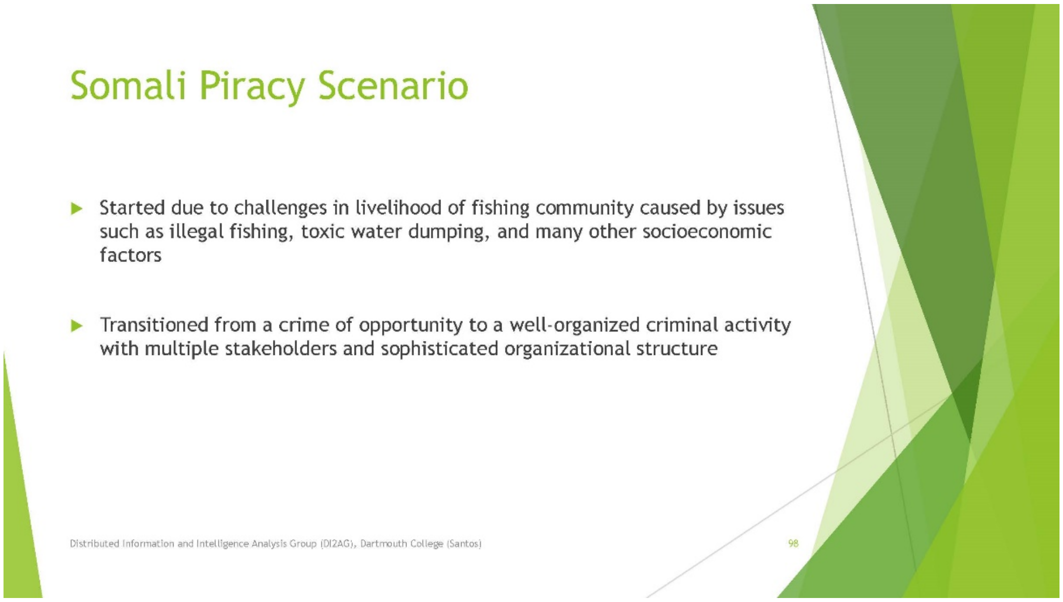
[*] Thayer School of Engineering, Hanover, NH 03784, eugene.santos.jr@dartmouth.edu
[†] Department of Computer Science, Chicago, IL 60616, eunice.santos@iit.edu
[‡] Department of Computer Science, Chicago, IL 60616, jkorah3@iit.edu



Motivation

- Intelligence, Surveillance, and Reconnaissance (ISR)
  - Produce timely, correct, and actionable intelligence for the warfighter
  - Realities include partial/limited observability, pop-up "learning" adversaries, and resource-constrained intelligence asset, etc.
- Challenges
  - Dynamic, uncertain, and fluid environment to operate in
  - Stress of time and tight windows of opportunities
  - Multi-entity, multi-level, and multi-scale situations involving multiple friends, foes, and neutrals
    - Islamic State (IS) in Iraq, Syria, eastern Libya and the Sinai Peninsula of Egypt
    - Outbreak of Ebola in Guinea, Sierra Leone and Liberia
  - Inherently convoluted and evolving nature of these ongoing events and their temporal urgency
- Goal
  - How to help commanders and decision-makers by modeling targets as complex adaptive systems?

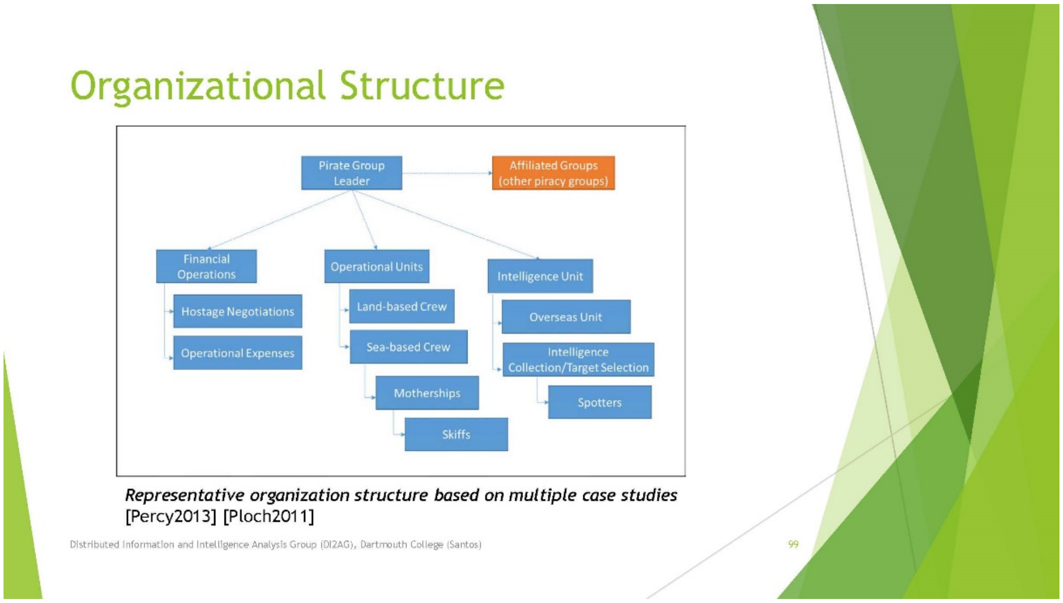
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 97



Somali Piracy Scenario

- Started due to challenges in livelihood of fishing community caused by issues such as illegal fishing, toxic water dumping, and many other socioeconomic factors
- Transitioned from a crime of opportunity to a well-organized criminal activity with multiple stakeholders and sophisticated organizational structure

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 98

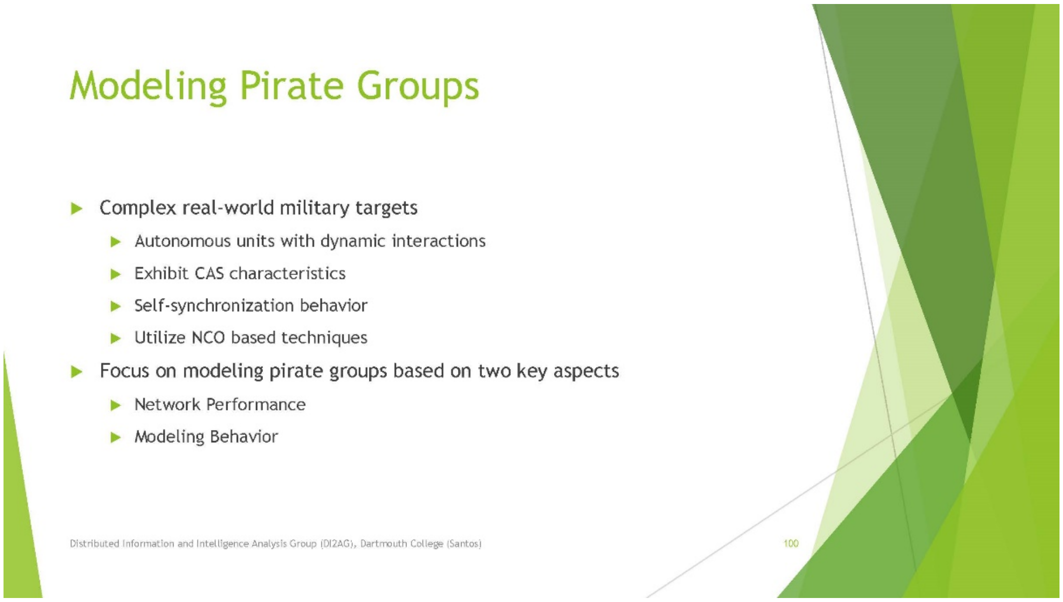


Organizational Structure

- Pirate Group Leader → Affiliated Groups (other piracy groups)
  - Financial Operations
    - Hostage Negotiations
    - Operational Expenses
  - Operational Units
    - Land-based Crew
    - Sea-based Crew
      - Motherships
        - Skiffs
  - Intelligence Unit
    - Overseas Unit
    - Intelligence Collection/Target Selection
      - Spotters

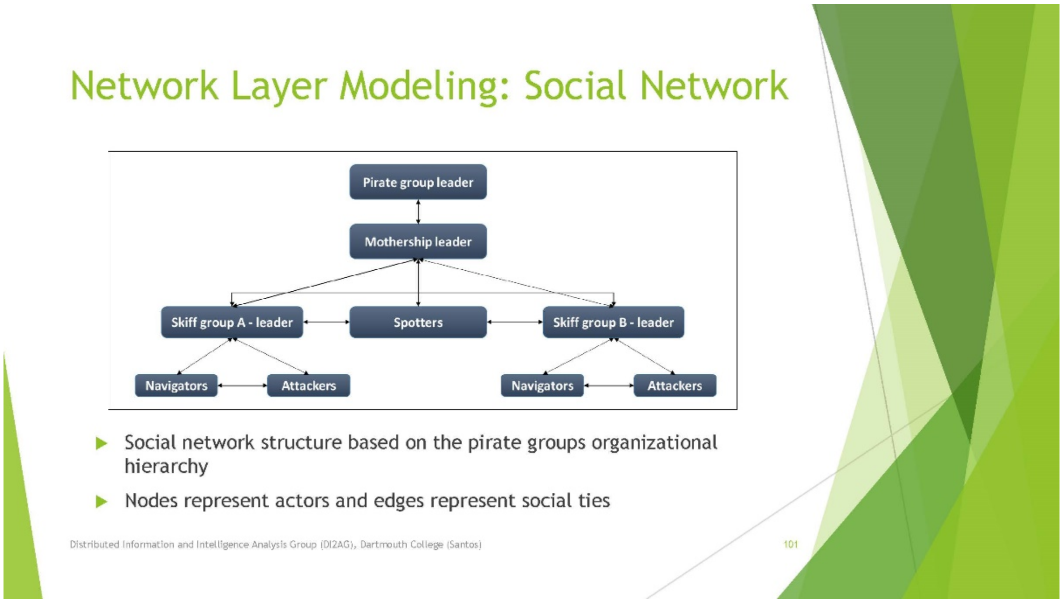
Representative organization structure based on multiple case studies
[Percy2013] [Ploch2011]

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 99



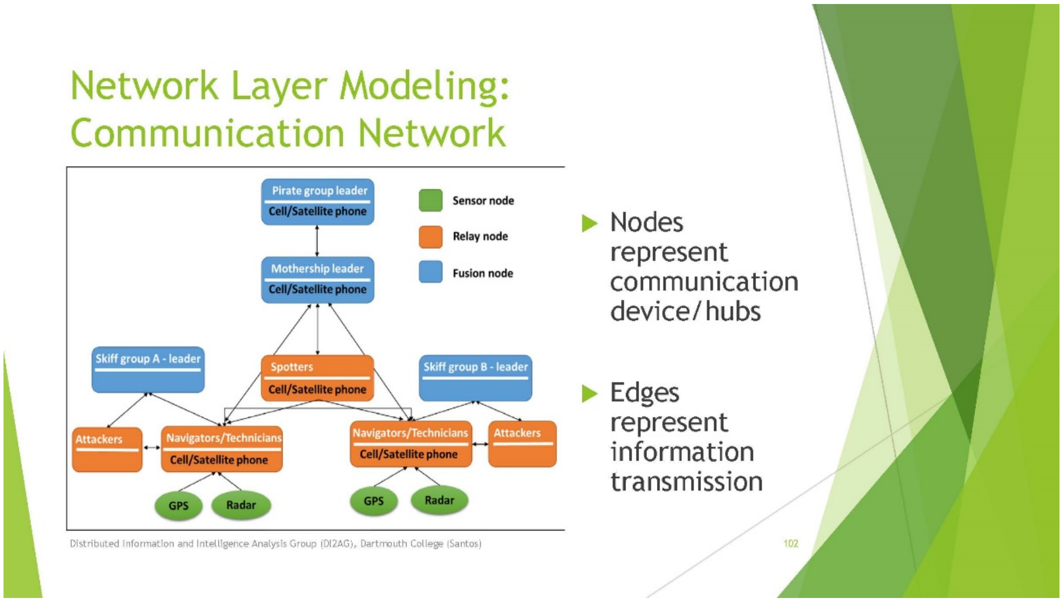
Modeling Pirate Groups

- Complex real-world military targets
  - Autonomous units with dynamic interactions
  - Exhibit CAS characteristics
  - Self-synchronization behavior
  - Utilize NCO based techniques
- Focus on modeling pirate groups based on two key aspects
  - Network Performance
  - Modeling Behavior



Network Layer Modeling: Social Network

- Social network structure based on the pirate groups organizational hierarchy
- Nodes represent actors and edges represent social ties



Network Layer Modeling: Communication Network

Pirate group leader
Cell/Satellite phone

Mothership leader
Cell/Satellite phone

Skiff group A - leader

Spotters
Cell/Satellite phone

Skiff group B - leader

Attackers

Navigators/Technicians
Cell/Satellite phone

Navigators/Technicians
Cell/Satellite phone

Attackers

GPS

Radar

GPS

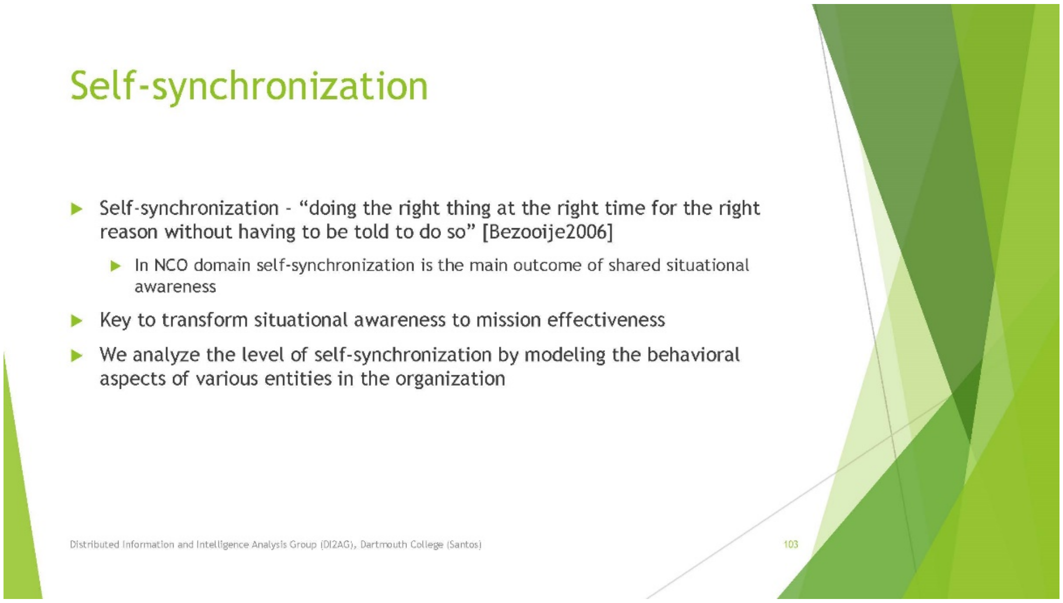
Radar

🟩 Sensor node

🟧 Relay node

🟦 Fusion node

- Nodes represent communication device/hubs
- Edges represent information transmission



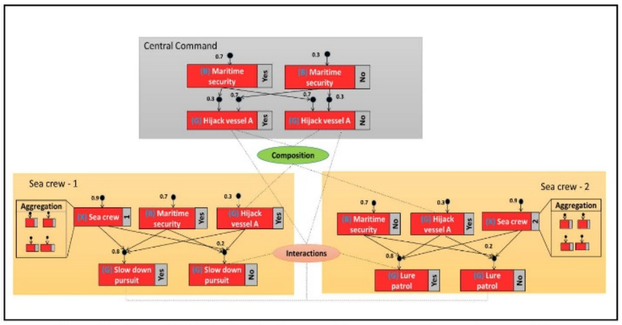
Self-synchronization

- Self-synchronization - “doing the right thing at the right time for the right reason without having to be told to do so” [Bezooije2006]
  - In NCO domain self-synchronization is the main outcome of shared situational awareness
- Key to transform situational awareness to mission effectiveness
- We analyze the level of self-synchronization by modeling the behavioral aspects of various entities in the organization

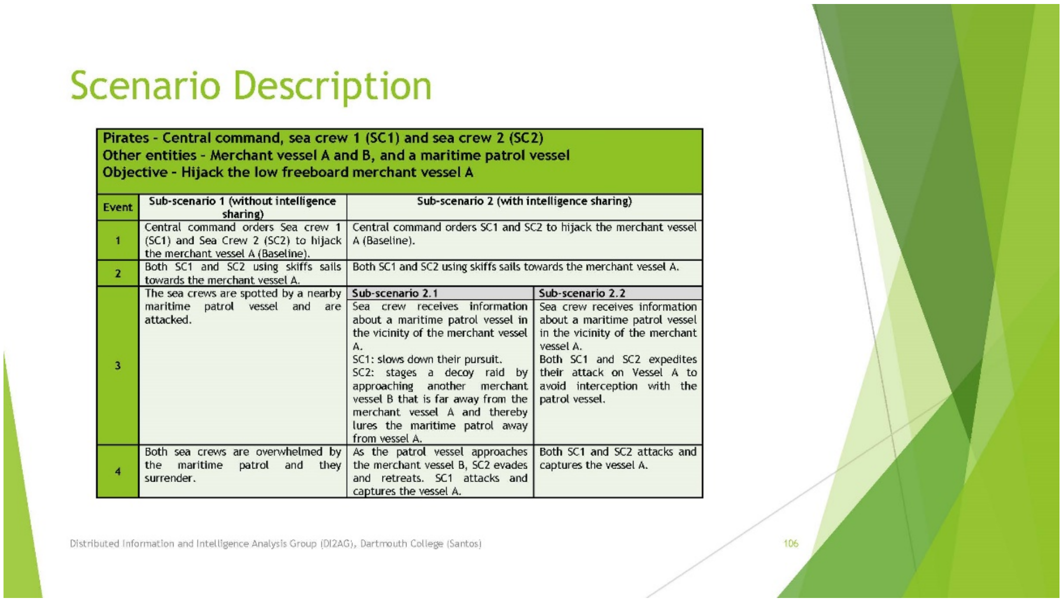
Information Aggregation

- Situational awareness is a critical aspect for effective self-synchronization in complex military systems
- We use caBKBs to model a pirate group's behavior based on situational awareness
  - Achieve information aggregation by combining multiple sources to model overall behavior
- Evaluate the effectiveness of self-synchronization based on situational awareness and actions performed by the entities

Information Aggregation cont.



Example of a caBKB illustrating information aggregation and composition



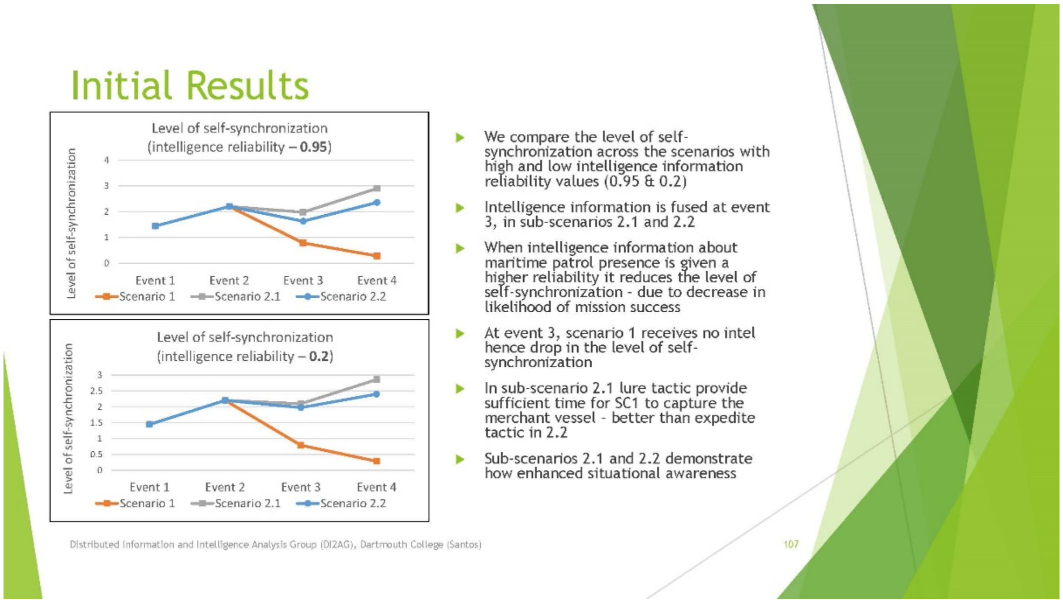
Scenario Description

Pirates - Central command, sea crew 1 (SC1) and sea crew 2 (SC2)
Other entities - Merchant vessel A and B, and a maritime patrol vessel
Objective - Hijack the low freeboard merchant vessel A

Event	Sub-scenario 1 (without intelligence sharing)	Sub-scenario 2 (with intelligence sharing)	
1	Central command orders Sea crew 1 (SC1) and Sea Crew 2 (SC2) to hijack the merchant vessel A (Baseline).	Central command orders SC1 and SC2 to hijack the merchant vessel A (Baseline).	
2	Both SC1 and SC2 using skiffs sails towards the merchant vessel A.	Both SC1 and SC2 using skiffs sails towards the merchant vessel A.	
3	The sea crews are spotted by a nearby maritime patrol vessel and are attacked.	**Sub-scenario 2.1** Sea crew receives information about a maritime patrol vessel in the vicinity of the merchant vessel A. SC1: slows down their pursuit. SC2: stages a decoy raid by approaching another merchant vessel B that is far away from the merchant vessel A and thereby lures the maritime patrol away from vessel A.	**Sub-scenario 2.2** Sea crew receives information about a maritime patrol vessel in the vicinity of the merchant vessel A. Both SC1 and SC2 expedites their attack on Vessel A to avoid interception with the patrol vessel.
4	Both sea crews are overwhelmed by the maritime patrol and they surrender.	As the patrol vessel approaches the merchant vessel B, SC2 evades and retreats. SC1 attacks and captures the vessel A.	Both SC1 and SC2 attacks and captures the vessel A.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

106

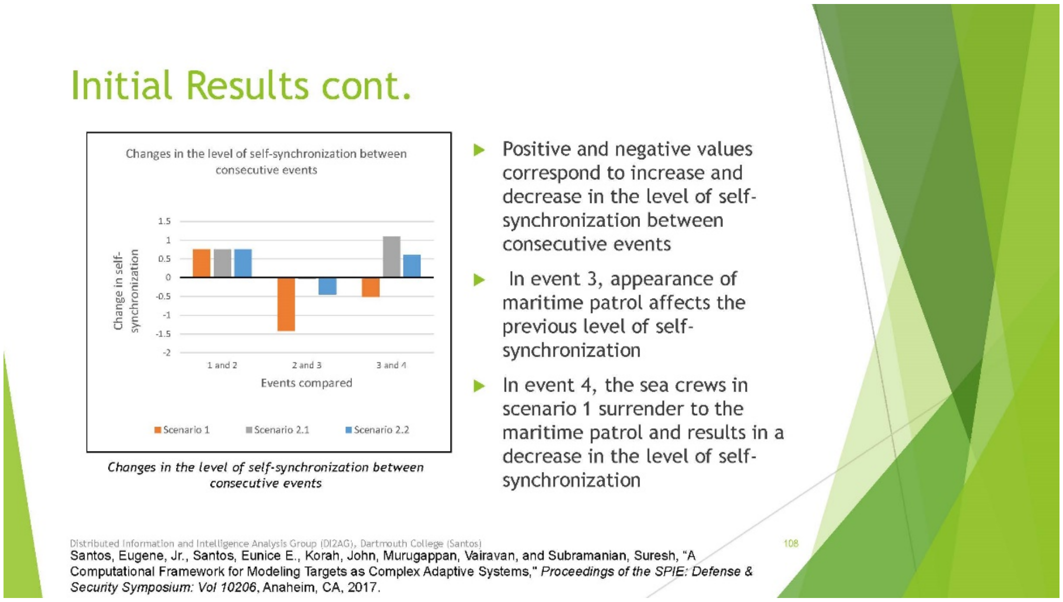


Initial Results

- We compare the level of self-synchronization across the scenarios with high and low intelligence information reliability values (0.95 & 0.2)
- Intelligence information is fused at event 3, in sub-scenarios 2.1 and 2.2
- When intelligence information about maritime patrol presence is given a higher reliability it reduces the level of self-synchronization - due to decrease in likelihood of mission success
- At event 3, scenario 1 receives no intel hence drop in the level of self-synchronization
- In sub-scenario 2.1 lure tactic provide sufficient time for SC1 to capture the merchant vessel - better than expedite tactic in 2.2
- Sub-scenarios 2.1 and 2.2 demonstrate how enhanced situational awareness

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos)

107



Initial Results cont.

Changes in the level of self-synchronization between consecutive events

- Positive and negative values correspond to increase and decrease in the level of self-synchronization between consecutive events
- In event 3, appearance of maritime patrol affects the previous level of self-synchronization
- In event 4, the sea crews in scenario 1 surrender to the maritime patrol and results in a decrease in the level of self-synchronization

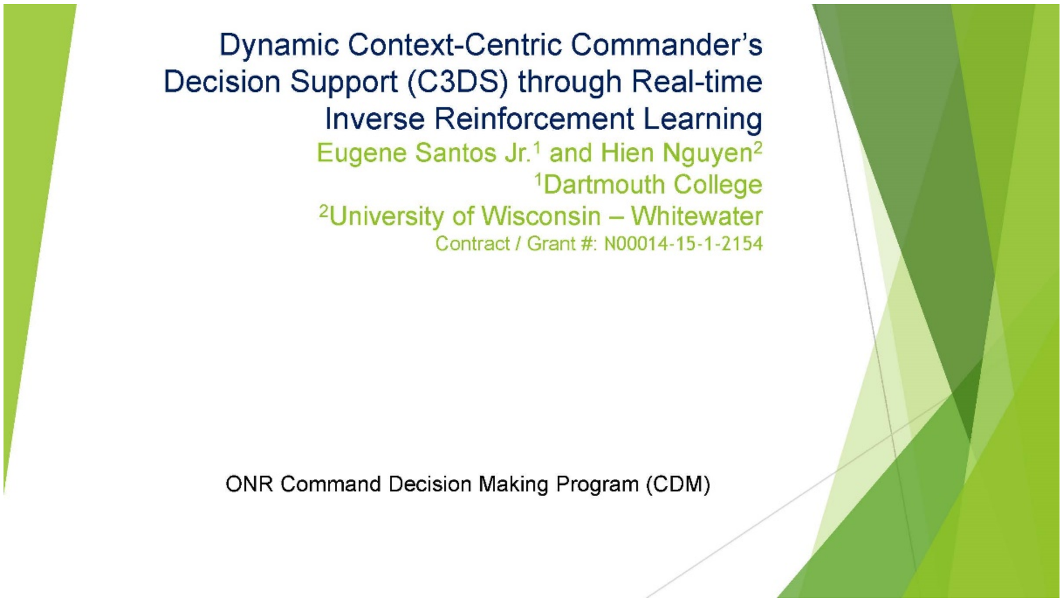
Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 108

Santos, Eugene, Jr., Santos, Eunice E., Korah, John, Murugappan, Vairavan, and Subramanian, Suresh, "A Computational Framework for Modeling Targets as Complex Adaptive Systems," *Proceedings of the SPIE: Defense & Security Symposium: Vol 10206*, Anaheim, CA, 2017.

Overview

- Adversary Intent Inferencing (AII)
- Adversarial Modeling
- Evolution of AII – Projects and Domains
  - Complexity, Computation, and Capabilities
  - Static Intent Model, Stochastic Behavior, Single Adversary
  - Static Intent Model, Dynamic Behavior, Single Adversary
  - Dynamic Intent Model, Evolving Behavior, Single Adversary
  - Networked Intent Model, Evolving Behaviors, Multiple Adversaries
- Future Work
  - Learning Adversary Intent and Decision Models

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 109



Dynamic Context-Centric Commander's Decision Support (C3DS) through Real-time Inverse Reinforcement Learning

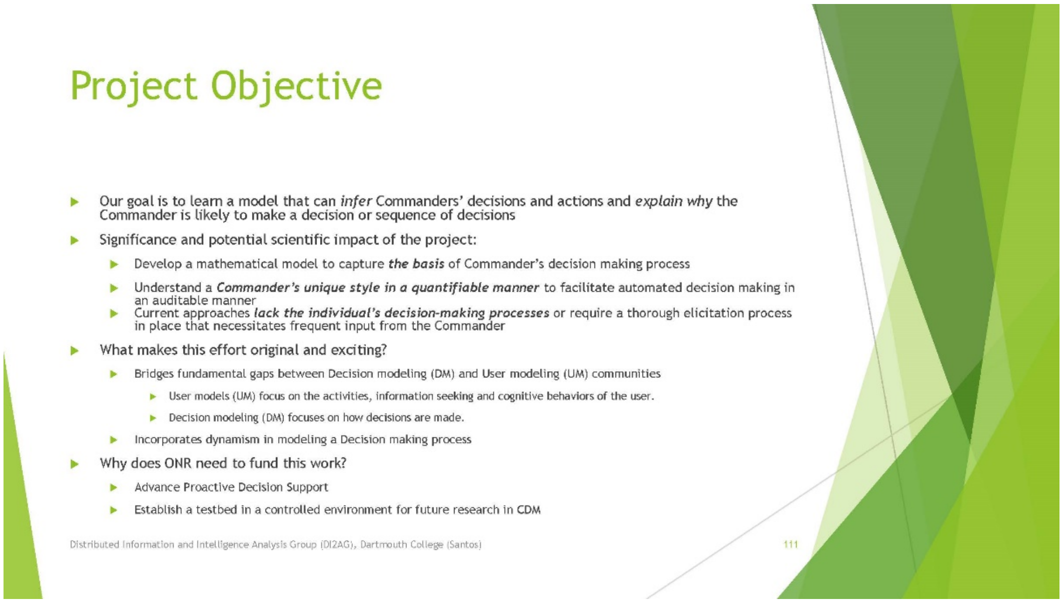
Eugene Santos Jr.[1] and Hien Nguyen[2]

[1]Dartmouth College

[2]University of Wisconsin – Whitewater

Contract / Grant #: N00014-15-1-2154

ONR Command Decision Making Program (CDM)



Project Objective

- Our goal is to learn a model that can *infer* Commanders' decisions and actions and *explain why* the Commander is likely to make a decision or sequence of decisions
- Significance and potential scientific impact of the project:
  - Develop a mathematical model to capture ***the basis*** of Commander's decision making process
  - Understand a ***Commander's unique style in a quantifiable manner*** to facilitate automated decision making in an auditable manner
  - Current approaches ***lack the individual's decision-making processes*** or require a thorough elicitation process in place that necessitates frequent input from the Commander
- What makes this effort original and exciting?
  - Bridges fundamental gaps between Decision modeling (DM) and User modeling (UM) communities
    - User models (UM) focus on the activities, information seeking and cognitive behaviors of the user.
    - Decision modeling (DM) focuses on how decisions are made.
  - Incorporates dynamism in modeling a Decision making process
- Why does ONR need to fund this work?
  - Advance Proactive Decision Support
  - Establish a testbed in a controlled environment for future research in CDM

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 111

Technical approach - Scenarios

- Commander 1's battle in Steel Ocean

- Commander 2's battle in Steel Ocean

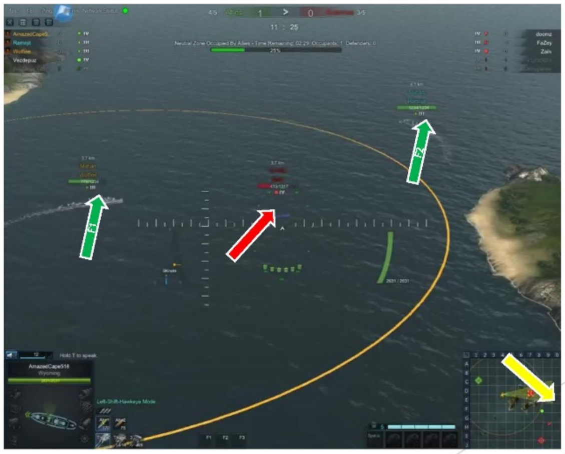
How to capture Commanders' decision making process?

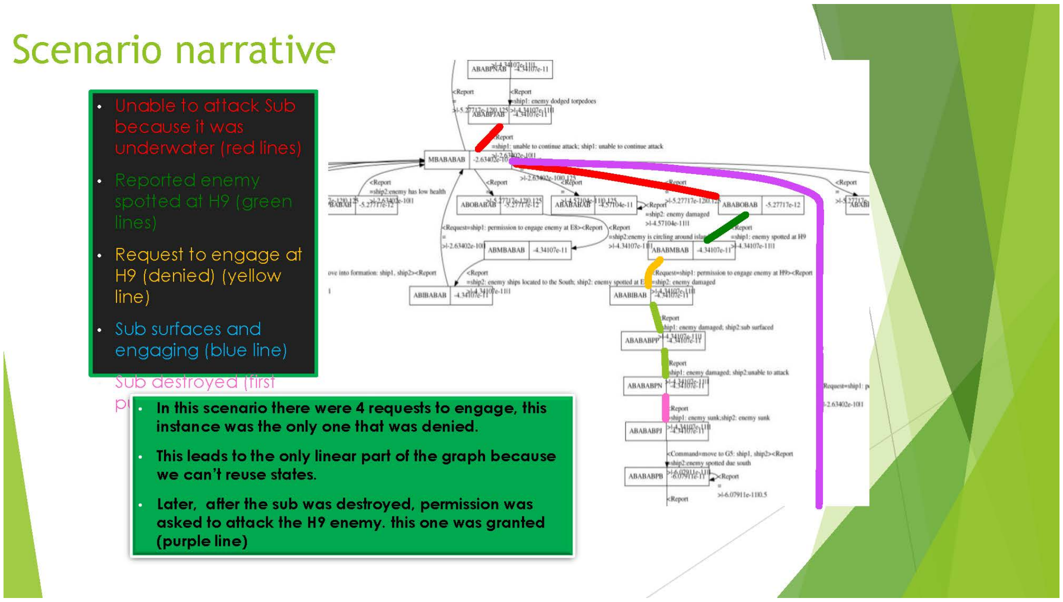
How to capture Commanders' decision making styles?

How to assess the quality of Commanders' decisions?

Scenario narrative

- Enemy spotted at H9
- Enemy submarine under water
- Friendly units (F1, F2)
- While submarine was hidden under water, F1 requested to engage enemy at H9
- This request was denied, and both F1 and F2 were told to focus the sub
- Ultimately, they destroyed the submarine with minimal interference from the enemy at H9





ABABPNAB
-4.34107e-11
<Report
=ship1: enemy dodged torpedoes
ABABPJAB
-4.34107e-11
<Report
=ship1: unable to continue attack; ship1: unable to continue attack
MBABABAB
-2.63402e-10
<Report
=ship2:enemy has low health
ABOBABAB
-5.27717e-12
ABABABAB
-4.57104e-11
<Report
=ship2: enemy damaged
ABABOBAB
-5.27717e-12
<Request=ship1: permission to engage enemy at E8><Report
ABMBABAB
-4.34107e-11
<Report
=ship2:enemy is circling around isla
ABABMBAB
-4.34107e-11
<Report
=ship1: enemy spotted at H9
ove into formation: ship1, ship2><Report
<Report
=ship2: enemy ships located to the South; ship2: enemy spotted at E
ABIBABAB
-4.34107e-11
<Request=ship1: permission to engage enemy at H9><Report
=ship2: enemy damaged
ABABIBAB
-4.34107e-11
Report
ship1: enemy damaged; ship2:sub surfaced
ABABABPP
-4.34107e-11
Report
ship1: enemy damaged; ship2:unable to attack
ABABABPN
-4.34107e-11
Report
ship1: enemy sunk;ship2: enemy sunk
ABABABPJ
-4.34107e-11
<Command=move to G5: ship1, ship2><Report
=ship2:enemy spotted due south
ABABABPB
-6.07911e-11
<Report
>l-6.07911e-11l0.5
Request=ship1: p
-2.63402e-10l1

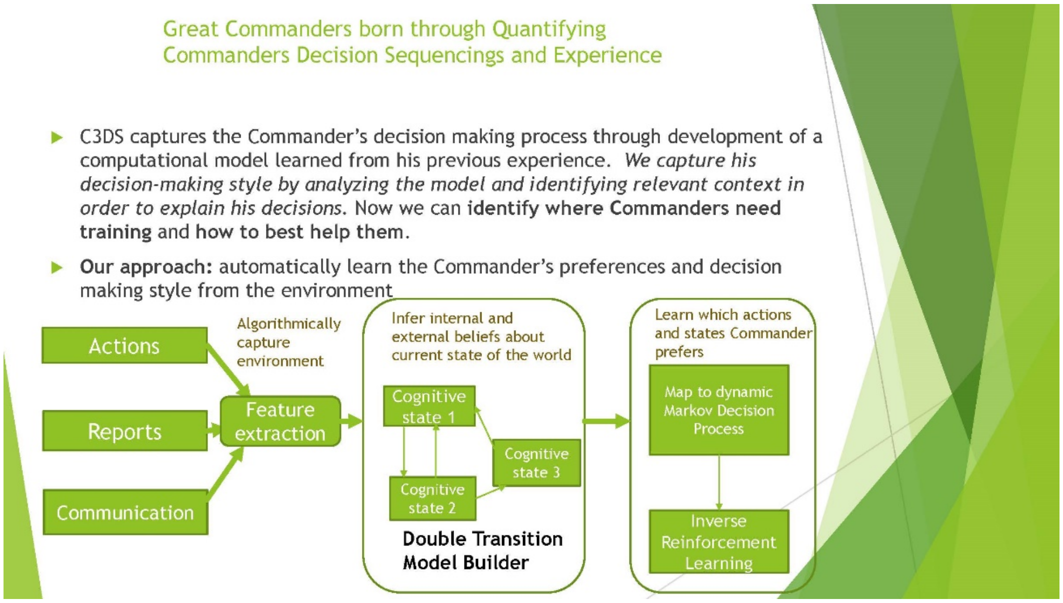
Scenario narrative

- Unable to attack Sub because it was underwater (red lines)
- Reported enemy spotted at H9 (green lines)
- Request to engage at H9 (denied) (yellow line)
- Sub surfaces and engaging (blue line)

Sub destroyed (first

p

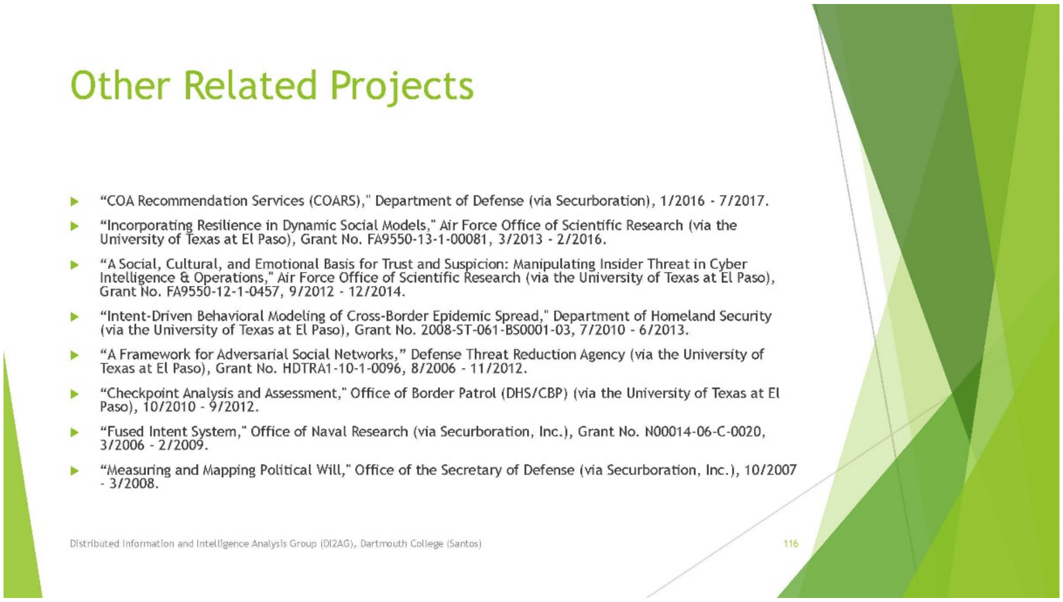
- In this scenario there were 4 requests to engage, this instance was the only one that was denied.
- This leads to the only linear part of the graph because we can't reuse states.
- Later, after the sub was destroyed, permission was asked to attack the H9 enemy. this one was granted (purple line)



Great Commanders born through Quantifying Commanders Decision Sequencings and Experience

- C3DS captures the Commander's decision making process through development of a computational model learned from his previous experience. *We capture his decision-making style by analyzing the model and identifying relevant context in order to explain his decisions.* Now we can **identify where Commanders need training** and **how to best help them**.
- **Our approach:** automatically learn the Commander's preferences and decision making style from the environment

Actions
Reports
Communication
Algorithmically capture environment
Feature extraction
Infer internal and external beliefs about current state of the world
Cognitive state 1
Cognitive state 2
Cognitive state 3
Double Transition Model Builder
Learn which actions and states Commander prefers
Map to dynamic Markov Decision Process
Inverse Reinforcement Learning



Other Related Projects

- "COA Recommendation Services (COARS)," Department of Defense (via Securboration), 1/2016 - 7/2017.
- "Incorporating Resilience in Dynamic Social Models," Air Force Office of Scientific Research (via the University of Texas at El Paso), Grant No. FA9550-13-1-00081, 3/2013 - 2/2016.
- "A Social, Cultural, and Emotional Basis for Trust and Suspicion: Manipulating Insider Threat in Cyber Intelligence & Operations," Air Force Office of Scientific Research (via the University of Texas at El Paso), Grant No. FA9550-12-1-0457, 9/2012 - 12/2014.
- "Intent-Driven Behavioral Modeling of Cross-Border Epidemic Spread," Department of Homeland Security (via the University of Texas at El Paso), Grant No. 2008-ST-061-BS0001-03, 7/2010 - 6/2013.
- "A Framework for Adversarial Social Networks," Defense Threat Reduction Agency (via the University of Texas at El Paso), Grant No. HDTRA1-10-1-0096, 8/2006 - 11/2012.
- "Checkpoint Analysis and Assessment," Office of Border Patrol (DHS/CBP) (via the University of Texas at El Paso), 10/2010 - 9/2012.
- "Fused Intent System," Office of Naval Research (via Securboration, Inc.), Grant No. N00014-06-C-0020, 3/2006 - 2/2009.
- "Measuring and Mapping Political Will," Office of the Secretary of Defense (via Securboration, Inc.), 10/2007 - 3/2008.

Distributed Information and Intelligence Analysis Group (DI2AG), Dartmouth College (Santos) 116

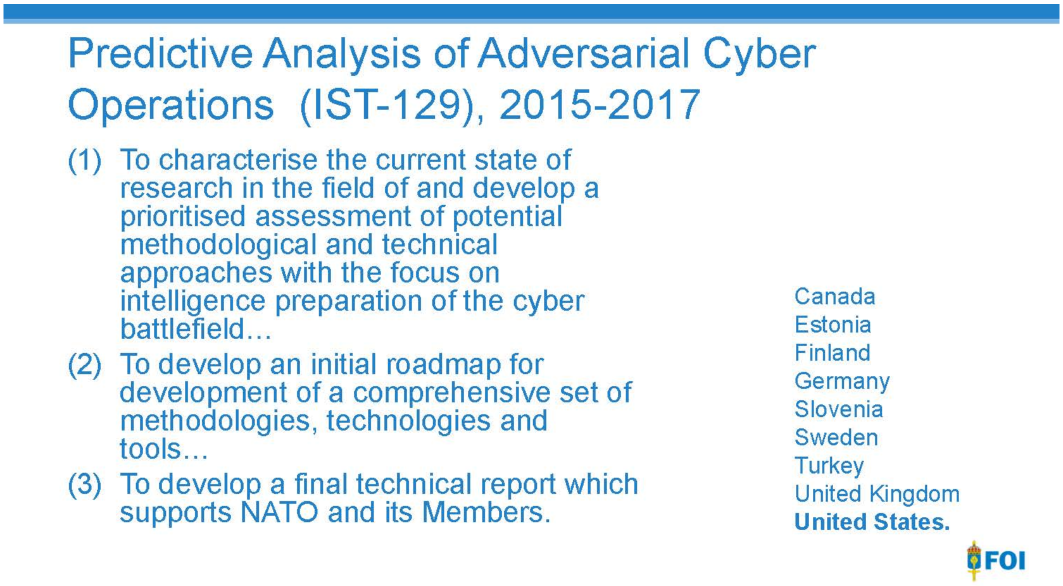
Position paper for IST-129

Predictive Analysis of Adversarial Cyber Operations

Teodor Sommestad
Swedish Defence Research Agency FOI

Foto: iStock Photo

FOI



Predictive Analysis of Adversarial Cyber Operations (IST-129), 2015-2017

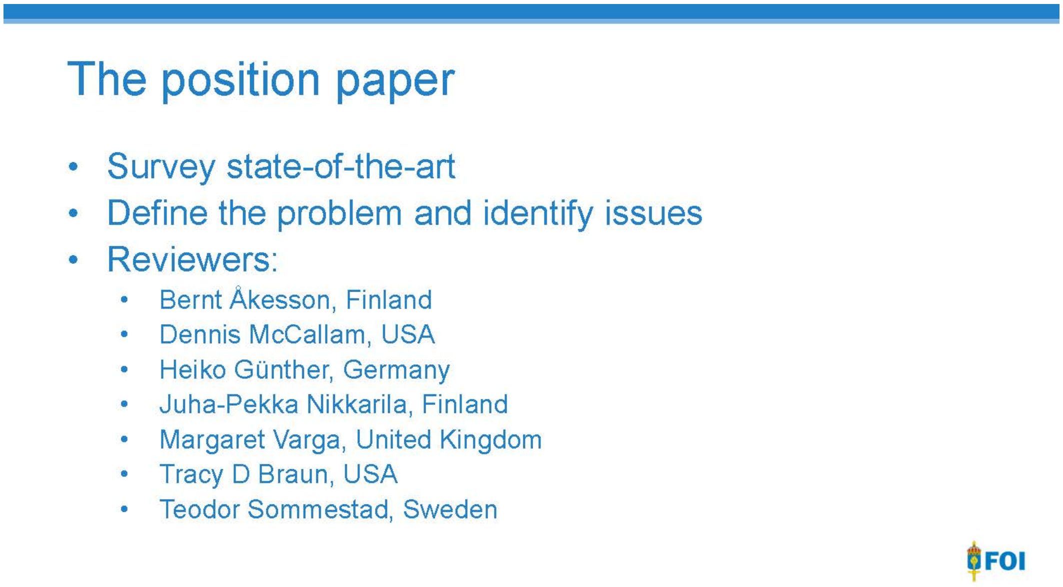
(1) To characterise the current state of research in the field of and develop a prioritised assessment of potential methodological and technical approaches with the focus on intelligence preparation of the cyber battlefield…

(2) To develop an initial roadmap for development of a comprehensive set of methodologies, technologies and tools…

(3) To develop a final technical report which supports NATO and its Members.

Canada
Estonia
Finland
Germany
Slovenia
Sweden
Turkey
United Kingdom
United States.

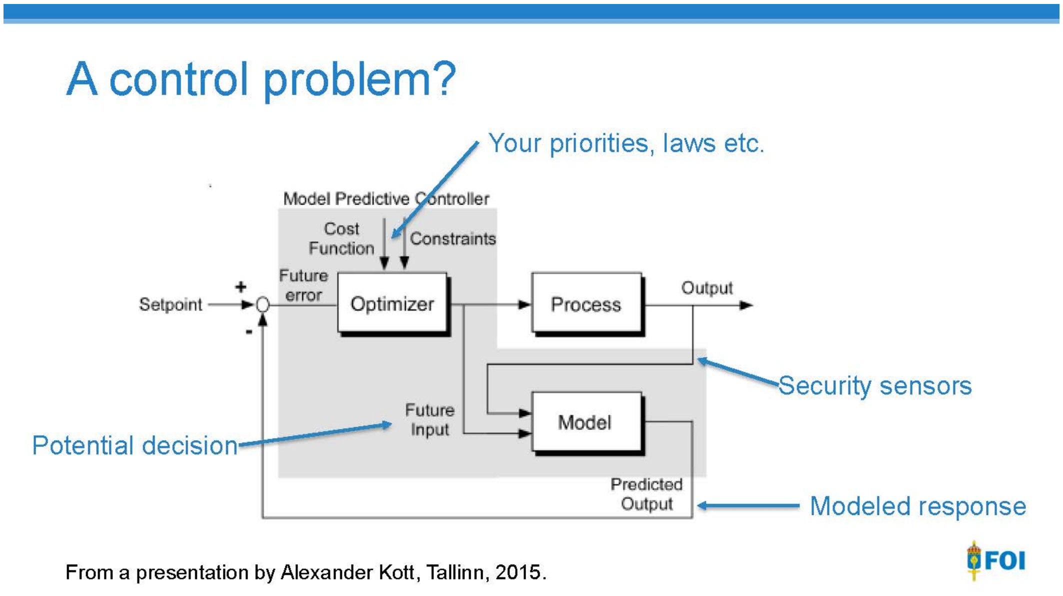
FOI



The position paper

- Survey state-of-the-art
- Define the problem and identify issues
- Reviewers:
  - Bernt Åkesson, Finland
  - Dennis McCallam, USA
  - Heiko Günther, Germany
  - Juha-Pekka Nikkarila, Finland
  - Margaret Varga, United Kingdom
  - Tracy D Braun, USA
  - Teodor Sommestad, Sweden

FOI



A control problem?

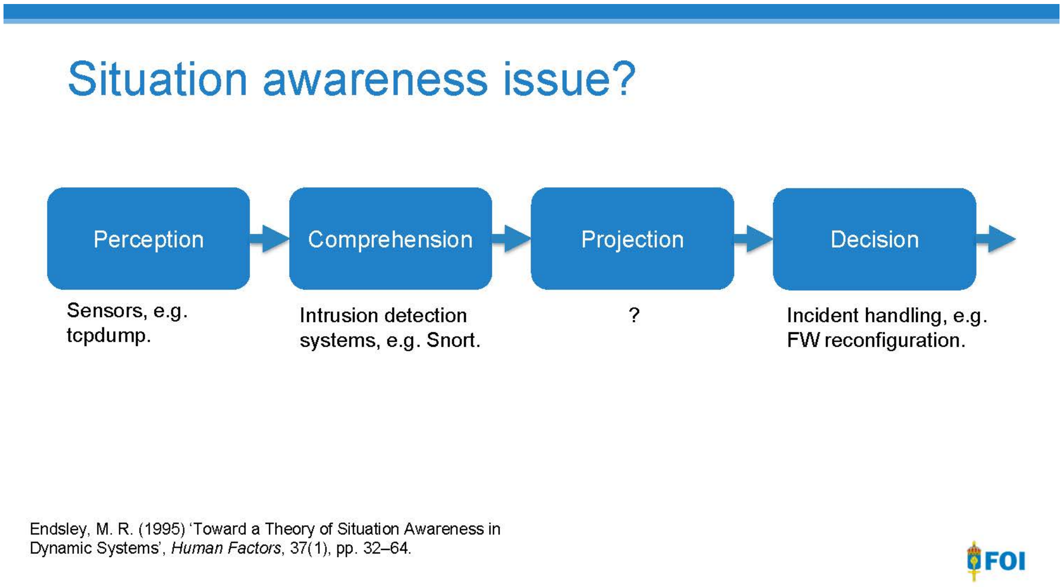
Your priorities, laws etc.

Security sensors

Potential decision

Modeled response

From a presentation by Alexander Kott, Tallinn, 2015.



Situation awareness issue?

Perception → Comprehension → Projection → Decision →

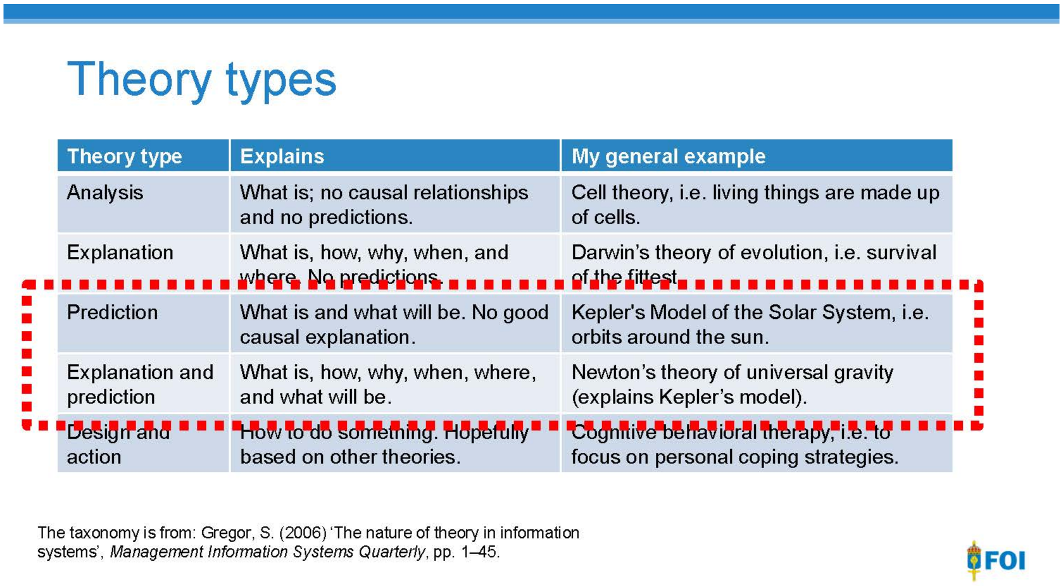
Perception: Sensors, e.g. tcpdump.

Comprehension: Intrusion detection systems, e.g. Snort.

Projection: ?

Decision: Incident handling, e.g. FW reconfiguration.

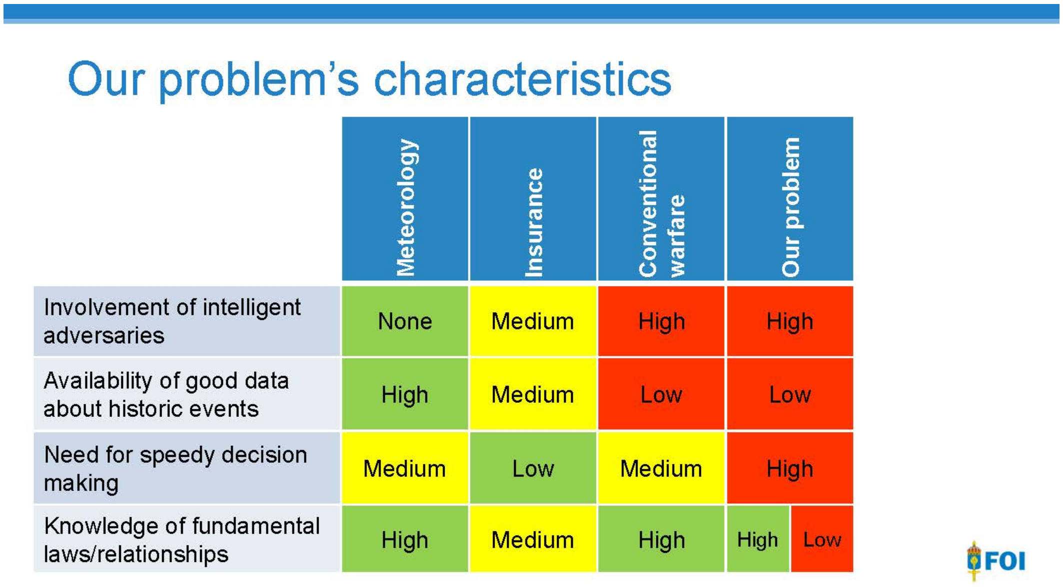
Endsley, M. R. (1995) 'Toward a Theory of Situation Awareness in Dynamic Systems', *Human Factors*, 37(1), pp. 32–64.



Theory types

Theory type	Explains	My general example
Analysis	What is; no causal relationships and no predictions.	Cell theory, i.e. living things are made up of cells.
Explanation	What is, how, why, when, and where. No predictions.	Darwin's theory of evolution, i.e. survival of the fittest.
Prediction	What is and what will be. No good causal explanation.	Kepler's Model of the Solar System, i.e. orbits around the sun.
Explanation and prediction	What is, how, why, when, where, and what will be.	Newton's theory of universal gravity (explains Kepler's model).
Design and action	How to do something. Hopefully based on other theories.	Cognitive behavioral therapy, i.e. to focus on personal coping strategies.

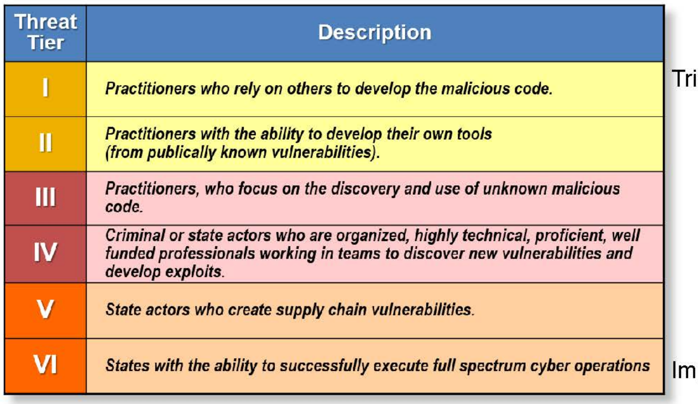
The taxonomy is from: Gregor, S. (2006) 'The nature of theory in information systems', *Management Information Systems Quarterly*, pp. 1–45.



Our problem's characteristics

	Meteorology	Insurance	Conventional warfare	Our problem
Involvement of intelligent adversaries	None	Medium	High	High
Availability of good data about historic events	High	Medium	Low	Low
Need for speedy decision making	Medium	Low	Medium	High
Knowledge of fundamental laws/relationships	High	Medium	High	High Low

The type of adversary



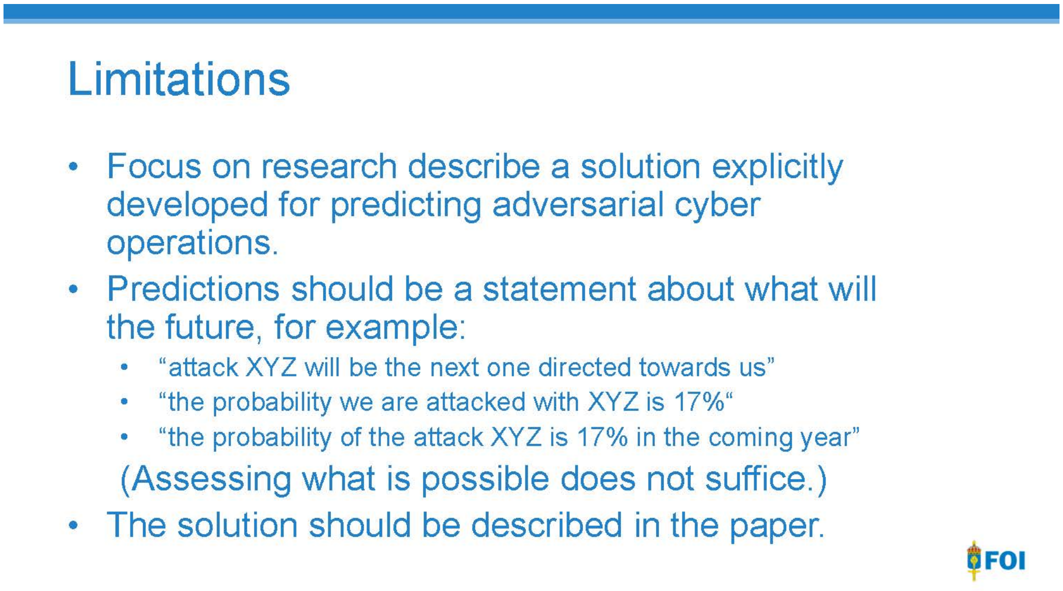
Threat Tier	Description
I	*Practitioners who rely on others to develop the malicious code.*
II	*Practitioners with the ability to develop their own tools (from publically known vulnerabilities).*
III	*Practitioners, who focus on the discovery and use of unknown malicious code.*
IV	*Criminal or state actors who are organized, highly technical, proficient, well funded professionals working in teams to discover new vulnerabilities and develop exploits.*
V	*State actors who create supply chain vulnerabilities.*
VI	*States with the ability to successfully execute full spectrum cyber operations*

Trivial: antivirus & firewalls

Impossible: uncertainty overflow

Defense Science Board Task Force Report: Resilient Military Systems and the Advanced Cyber Threat , January 2013

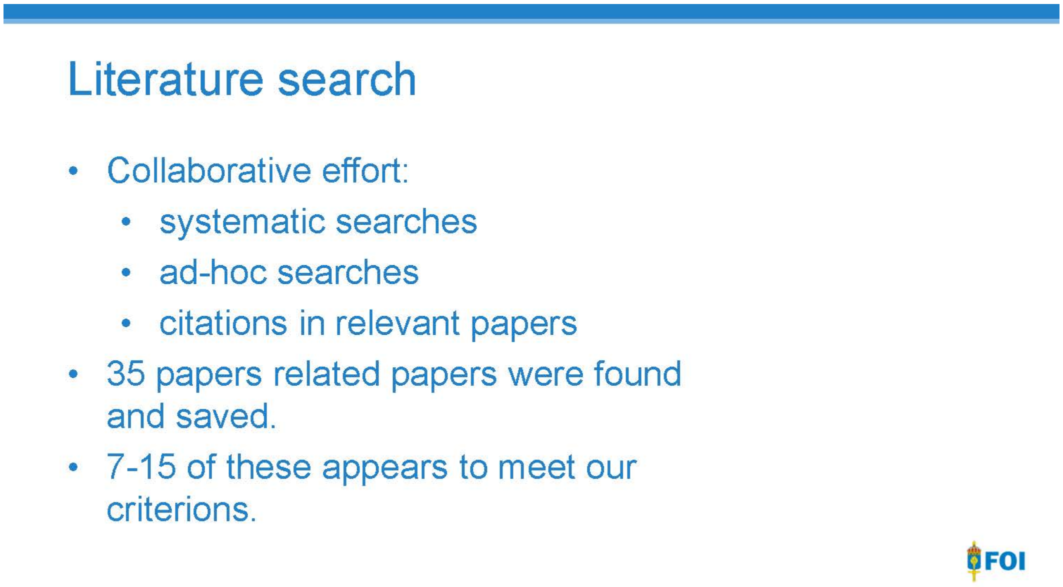
Our review



Limitations

- Focus on research describe a solution explicitly developed for predicting adversarial cyber operations.
- Predictions should be a statement about what will the future, for example:
  - “attack XYZ will be the next one directed towards us”
  - “the probability we are attacked with XYZ is 17%“
  - “the probability of the attack XYZ is 17% in the coming year”

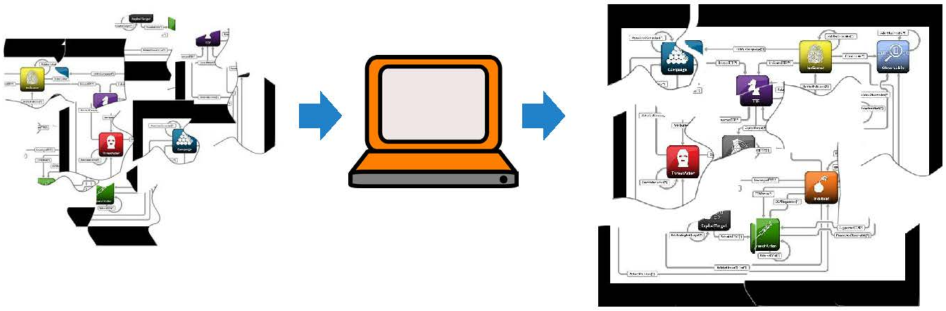
  (Assessing what is possible does not suffice.)
- The solution should be described in the paper.



Literature search

- Collaborative effort:
  - systematic searches
  - ad-hoc searches
  - citations in relevant papers
- 35 papers related papers were found and saved.
- 7-15 of these appears to meet our criterions.

Planned for an input-output-perspective

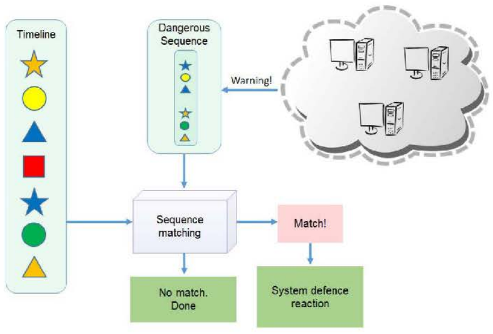


Three types of papers

1. General (conceptual) ideas on what to do
2. Extensions or applications of attack graphs
3. Concrete proposals

1. General idea on what to do (example)

1. Look for a dangerous sequence in your network.
2. If you see something dangerous, share the info with others.
3. The other should look for this pattern in their network; “an attack can be in earlier stages.”



V. Degeler, R. French, and K. Jones, “Self-Healing Intrusion Detection System Concept,” Proc. - 2nd IEEE Int. Conf. Big Data Secur. Cloud, IEEE BigDataSecurity 2016, 2nd IEEE Int. Conf. High Perform. Smart Comput. IEEE HPSC 2016 IEEE Int. Conf. Intell. Data Secur. IEEE IDS 2016, pp. 351–356, 2016

2. Attack graph based (example)

1. Build an attack graph from IDS alerts on some traffic (e.g. a honeypot).
2. Use the observed sequences to make predictions in future attacks.

“Predictability scores”, i.e. if it is expected based on what we seen before.

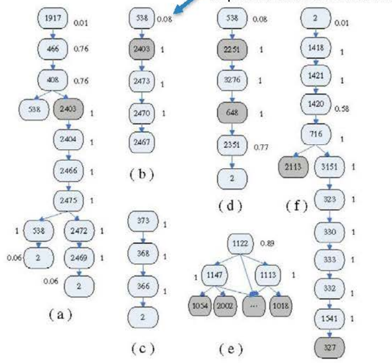
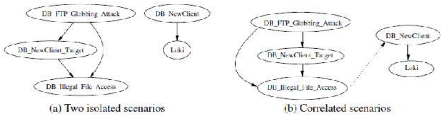


Figure 3. The Generated Attack Graph

J. Lei and Z. T. Li, “Using network attack graph to predict the future attacks,” Proc.Second Int. Conf. Commun. Netw. China, ChinaCom 2007, pp. 403–407, 2008.

2. Attack graph based (example 2)

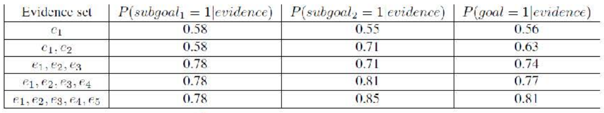
1. Have a forest of attack trees representing everything you care about.
2. Connect observed trees with each other in a Bayesian network.
3. Compute probabilities for different goals to be reached.



(a) Two isolated scenarios

(b) Correlated scenarios

Figure 3. Correlation of isolated scenarios



Evidence set	$P(subgoal_1 = 1 \mid evidence)$	$P(subgoal_2 = 1 \mid evidence)$	$P(goal = 1 \mid evidence)$
e_1	0.58	0.55	0.56
e_1, e_2	0.58	0.71	0.63
e_1, e_2, e_3	0.78	0.71	0.74
e_1, e_2, e_3, e_4	0.78	0.81	0.77
e_1, e_2, e_3, e_4, e_5	0.78	0.85	0.81

Is it expected based on what we seen before (i.e. $e_1 \ldots e_5$)?

X. Qin and W. Lee, “Attack plan recognition and prediction using causal networks,” Proc. - Annu.Comput. Secur. Appl. Conf. ACSAC, pp. 370–379, 2004.

3. Concrete proposals (example)

1. Look at historic events (in this case Spam).
2. Use machine learning and game theory to figure it out how it evolves over time.
3. Guess how future spam will look like.

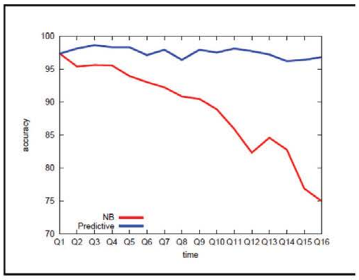


Figure 2. Results for the predictive defense case study. The plot shows how Spam filter accuracy (vertical axis) varies with time (horizontal axis) for the gold-standard NB filter (red) and Algorithm PD filter (blue).

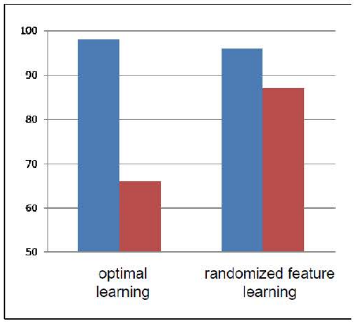
Colbaugh, R. and Glass, K. (2012) ‘Predictive defense against evolving adversaries’, *ISI 2012 - 2012 IEEE International Conference on Intelligence and Security Informatics: Cyberspace, Border, and Immigration Securities*, pp. 18–23.

Qualitative aspects

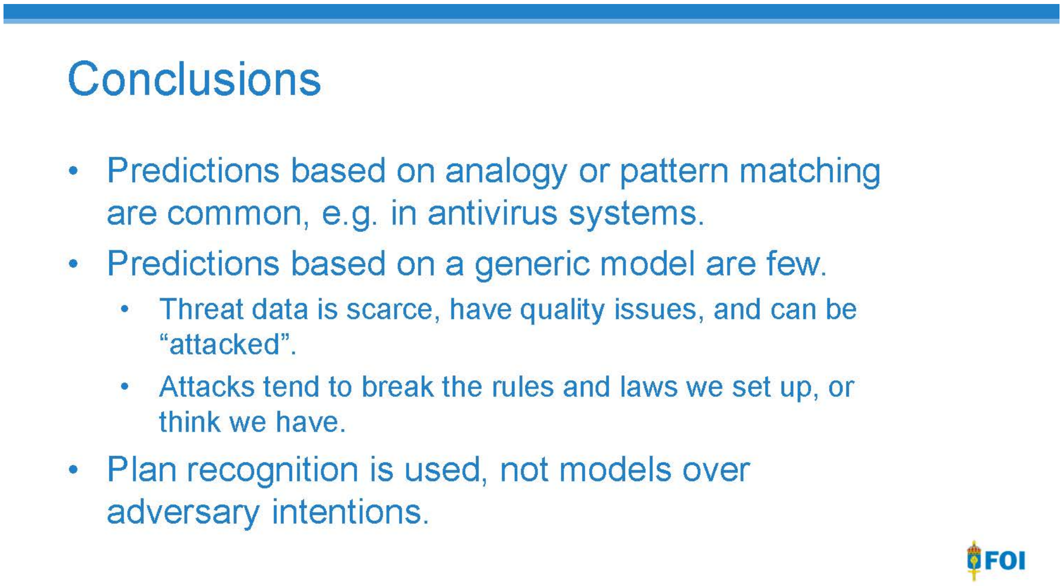
Issue	Treatment in the papers
Prediction accuracy and realism of tests	Some papers use realistic data and most discuss this.
Timing and decision support offered	Some papers discuss this and use it as an argument for predictions.
Where to find attack data, probabilities etc.	Many papers depend on this, but few address the problem.
Where to find data on the own network	Often addressed indirectly, e.g. as attack graphs are used.
Tampering with the data/algorithm used	Only a few papers address this at all.

Excerpts from the papers

- “There should be a known (or a highly probable) danger for the response to be triggered.” // Degeler et al.
- “In our approach, one of the most important components is the library of attack plans (defined as attack trees).” // X. Qin and W. Lee.
- “The approach has been tested using the real attack data collected in a honeynet.” //J. Lei and Z. T. Li,



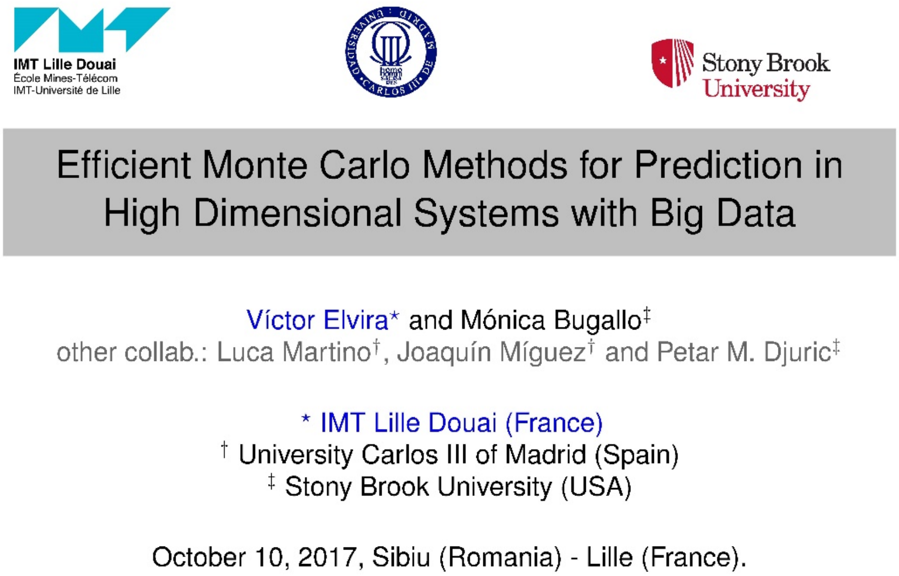
Only 66 % of smart attacks are detected with normal learning; 87% with a randomization on features to look for.



Conclusions

- Predictions based on analogy or pattern matching are common, e.g. in antivirus systems.
- Predictions based on a generic model are few.
  - Threat data is scarce, have quality issues, and can be "attacked".
  - Attacks tend to break the rules and laws we set up, or think we have.
- Plan recognition is used, not models over adversary intentions.

FOI



IMT Lille Douai
École Mines-Télécom
IMT-Université de Lille

Stony Brook University

Efficient Monte Carlo Methods for Prediction in High Dimensional Systems with Big Data

Víctor Elvira[*] and Mónica Bugallo[‡]
other collab.: Luca Martino[†], Joaquín Míguez[†] and Petar M. Djuric[‡]

[*] IMT Lille Douai (France)
[†] University Carlos III of Madrid (Spain)
[‡] Stony Brook University (USA)

October 10, 2017, Sibiu (Romania) - Lille (France).

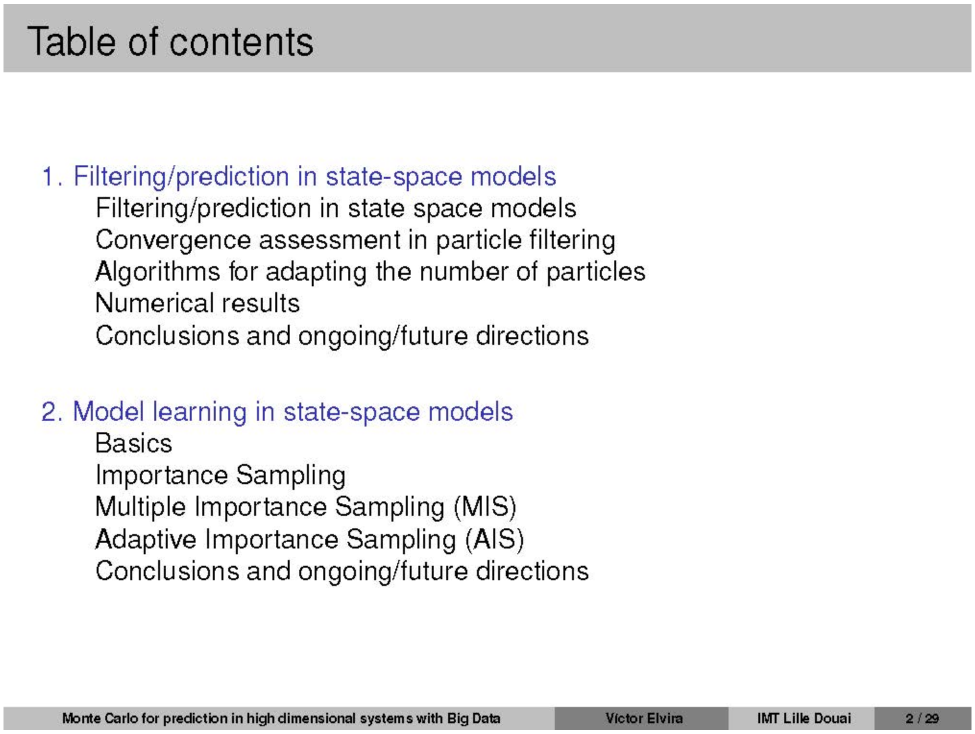
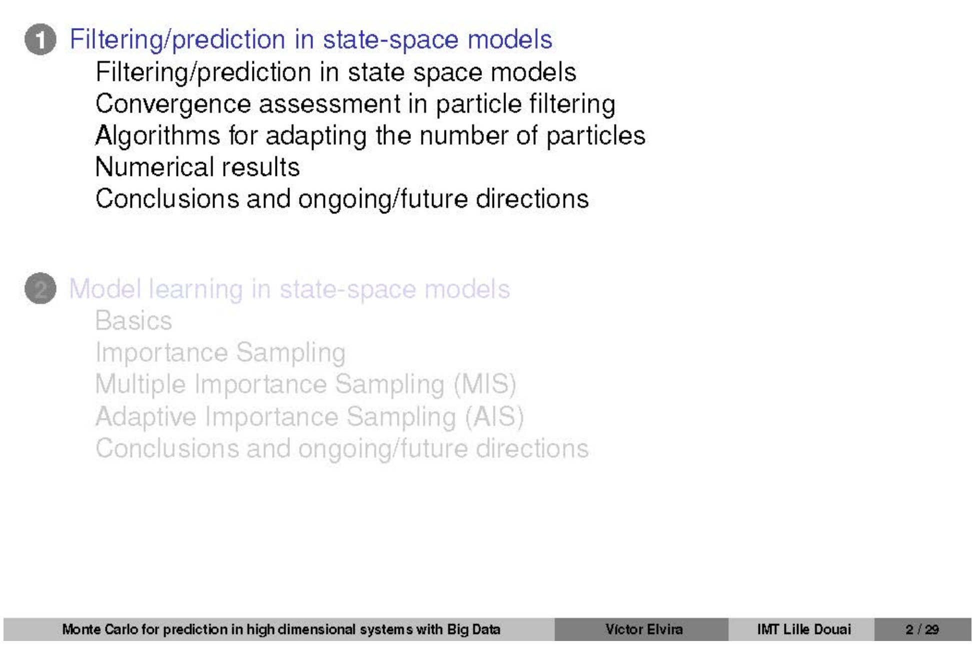


Table of contents

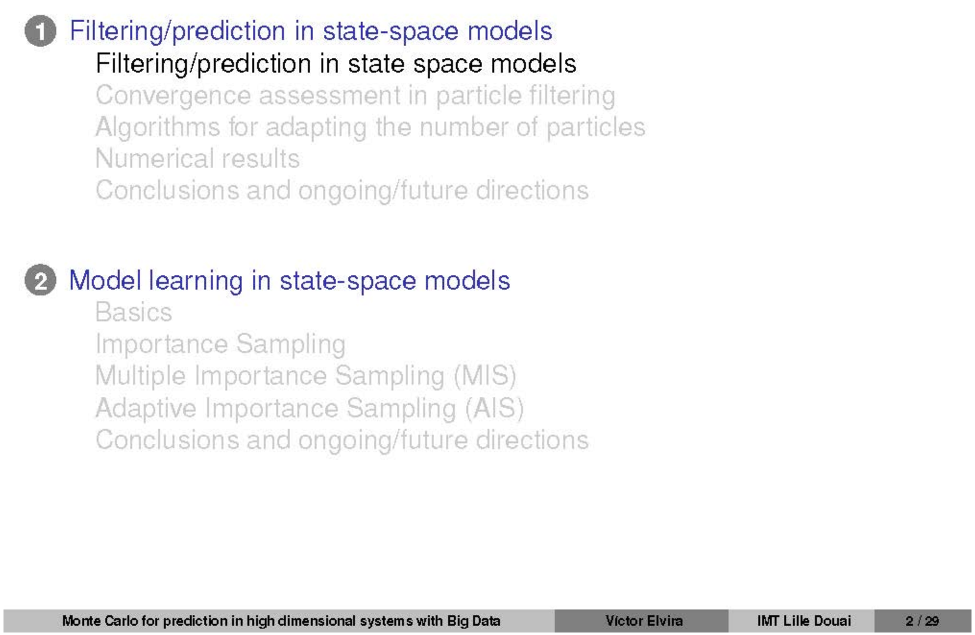
1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



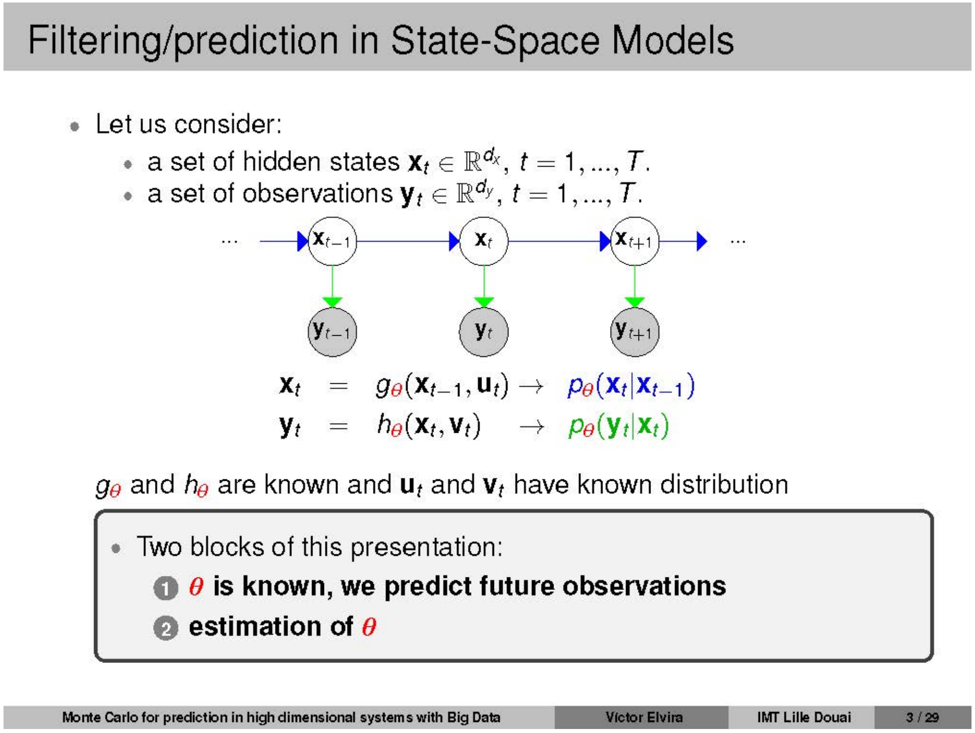
1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



Filtering/prediction in State-Space Models

- Let us consider:
  - a set of hidden states $\mathbf{x}_t \in \mathbb{R}^{d_x}$, $t = 1, ..., T$.
  - a set of observations $\mathbf{y}_t \in \mathbb{R}^{d_y}$, $t = 1, ..., T$.

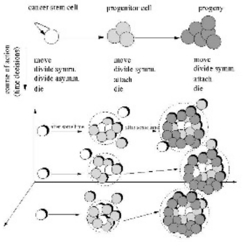
...
$\mathbf{x}_{t-1}$
$\mathbf{x}_t$
$\mathbf{x}_{t+1}$
...
$\mathbf{y}_{t-1}$
$\mathbf{y}_t$
$\mathbf{y}_{t+1}$

$$\begin{aligned}\mathbf{x}_t &= g_\theta(\mathbf{x}_{t-1}, \mathbf{u}_t) \rightarrow p_\theta(\mathbf{x}_t|\mathbf{x}_{t-1})\\ \mathbf{y}_t &= h_\theta(\mathbf{x}_t, \mathbf{v}_t) \quad \rightarrow p_\theta(\mathbf{y}_t|\mathbf{x}_t)\end{aligned}$$

g_θ and h_θ are known and $\mathbf{u}_t$ and $\mathbf{v}_t$ have known distribution

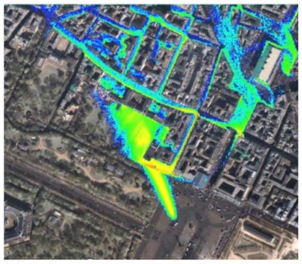
- Two blocks of this presentation:
  1. **θ is known, we predict future observations**
  2. **estimation of θ**

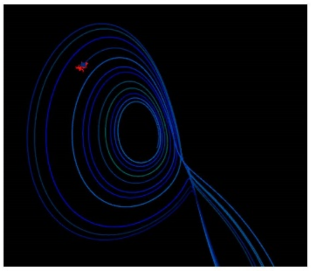
Examples of Dynamical Models



(a) Biology: cancer stem cells.

(b) Image tracking [link].



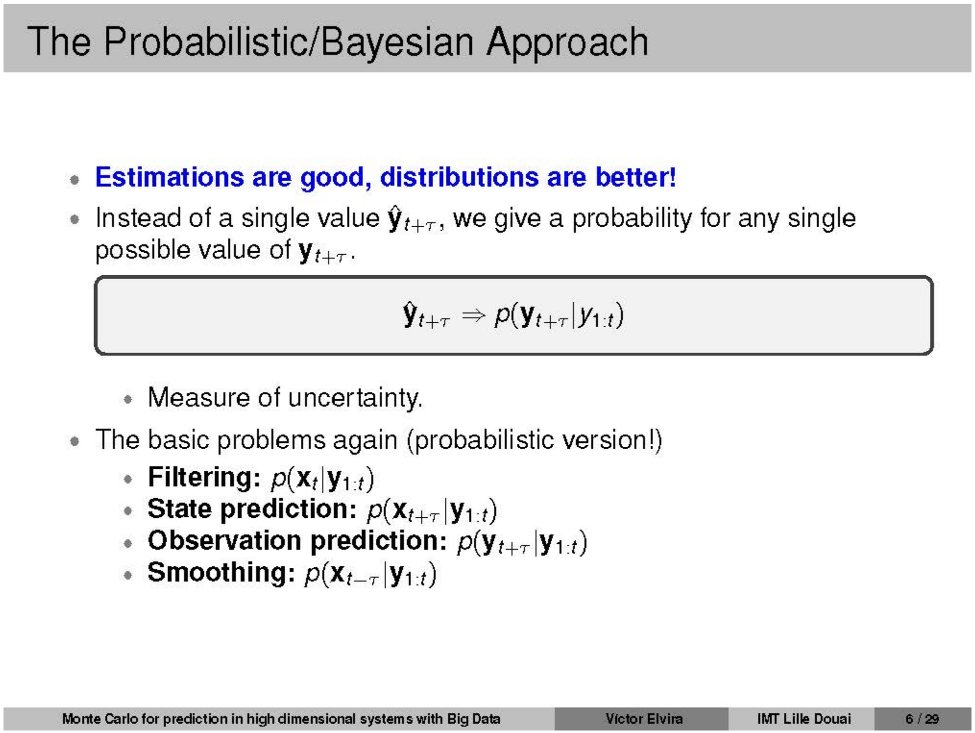


(c) Prediction of the propagation of the hazardous material.

(d) Chaotic systems, atmosphere: Lorenz63 model [link].

The inference/prediction problem

- We sequentially receive observations $\mathbf{y}_t$ related to the hidden state $\mathbf{x}_t$.
- At time t, we have accumulated t observations, $\mathbf{y}_{1:t} \equiv \{\mathbf{y}_1, ..., \mathbf{y}_t\}$.
- Interesting problems:
  - **Filtering**: estimate current state $\hat{\mathbf{x}}_t$ given $\mathbf{y}_{1:t}$
  - **Predict future state** $\hat{\mathbf{x}}_{t+\tau}$ given $\mathbf{y}_{1:t}$, $\tau \in \mathbb{N}^+$
  - **Predict future observation** $\hat{\mathbf{y}}_{t+\tau}$ given $\mathbf{y}_{1:t}$, $\tau \in \mathbb{N}^+$
  - **Smoothing**: estimate past state $\hat{\mathbf{x}}_{t-\tau}$ given $\mathbf{y}_{1:t}$, $\tau \in \mathbb{N}^+$
- We want to do it sequentially and efficiently.
  - At time t, we want to *process* only $\mathbf{y}_t$, but not reprocess all $\mathbf{y}_{1:t-1}$ (that were already processed!)

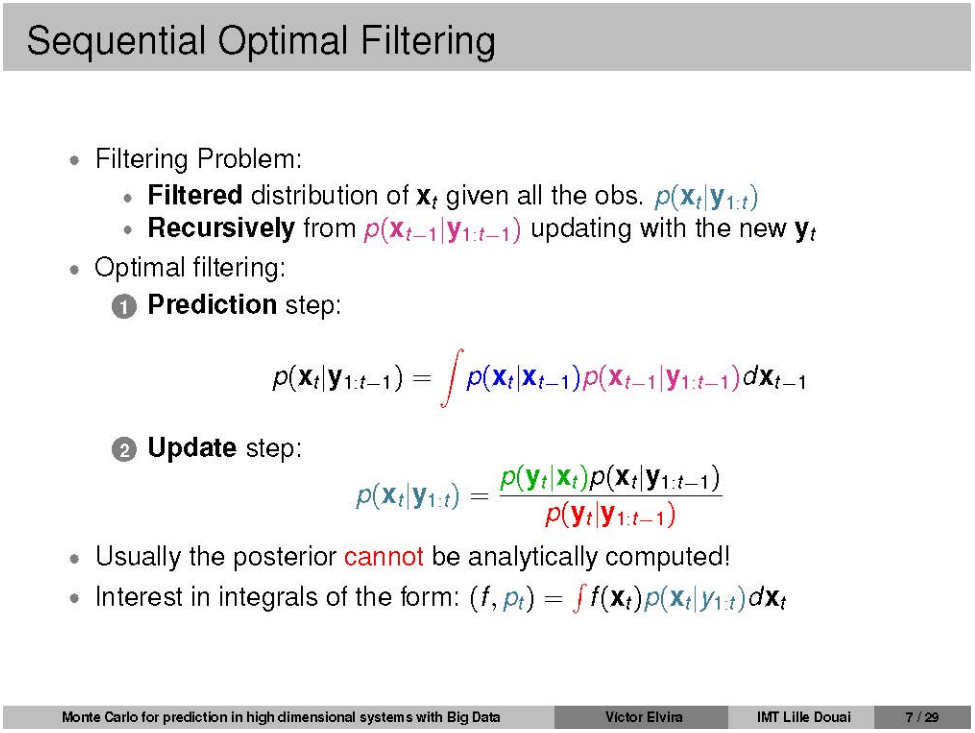


The Probabilistic/Bayesian Approach

- **Estimations are good, distributions are better!**
- Instead of a single value $\hat{\mathbf{y}}_{t+\tau}$, we give a probability for any single possible value of $\mathbf{y}_{t+\tau}$.

$$\hat{\mathbf{y}}_{t+\tau} \Rightarrow p(\mathbf{y}_{t+\tau}|y_{1:t})$$

  - Measure of uncertainty.
- The basic problems again (probabilistic version!)
  - **Filtering:** $p(\mathbf{x}_t|\mathbf{y}_{1:t})$
  - **State prediction:** $p(\mathbf{x}_{t+\tau}|\mathbf{y}_{1:t})$
  - **Observation prediction:** $p(\mathbf{y}_{t+\tau}|\mathbf{y}_{1:t})$
  - **Smoothing:** $p(\mathbf{x}_{t-\tau}|\mathbf{y}_{1:t})$



Sequential Optimal Filtering

- Filtering Problem:
  - **Filtered** distribution of $\mathbf{x}_t$ given all the obs. $p(\mathbf{x}_t|\mathbf{y}_{1:t})$
  - **Recursively** from $p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$ updating with the new $\mathbf{y}_t$
- Optimal filtering:
  1. **Prediction** step:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})d\mathbf{x}_{t-1}$$

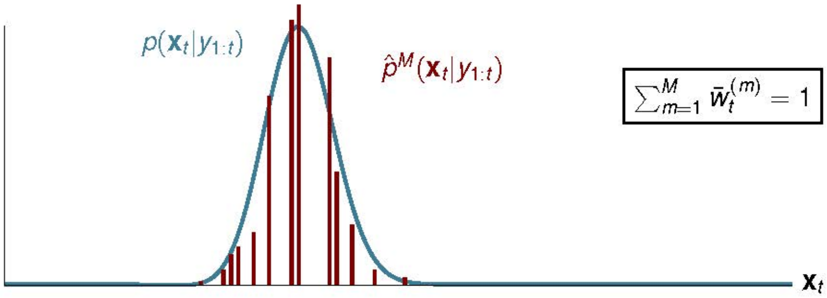
  2. **Update** step:

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}$$

- Usually the posterior cannot be analytically computed!
- Interest in integrals of the form: $(f, p_t) = \int f(\mathbf{x}_t)p(\mathbf{x}_t|y_{1:t})d\mathbf{x}_t$

Particle Filtering (Sequential Monte Carlo)

- The distributions are approximated by a random measure of M particles and associated normalized weights $\mathcal{X} = \{\mathbf{x}_t^{(m)}, \bar{w}_t^{(m)}\}_{m=1}^{M}$
  - $p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \hat{p}^M(\mathbf{x}_t|\mathbf{y}_{1:t}) = \sum_{m=1}^{M} \bar{w}_t^{(m)} \delta(\mathbf{x}_t - \mathbf{x}_t^{(m)})$

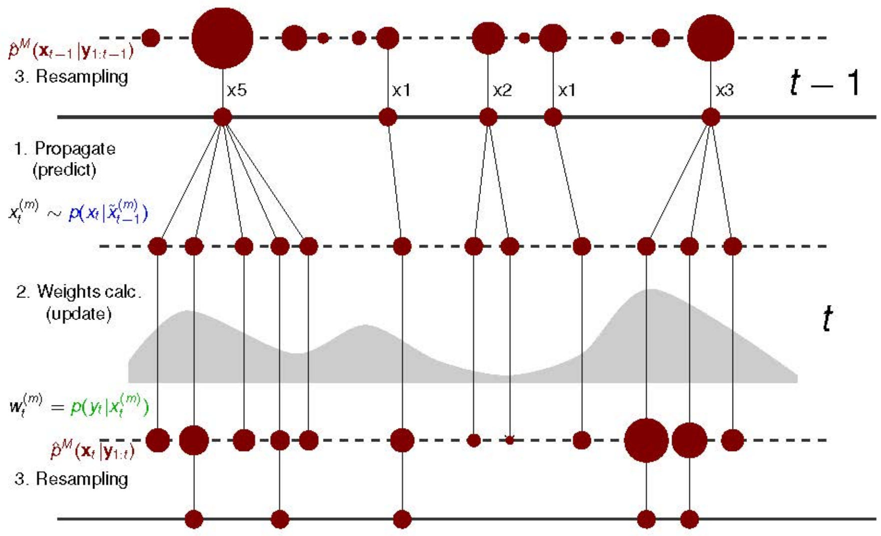


A Basic Particle Filter in a Nutshell

- Bootstrap PF $\equiv$ Sequential Importance Resampling (based on importance sampling) [Gordon93]
- The filtered distribution at time t is recursively approximated from $\tilde{\mathcal{X}}_{t-1} = \{\tilde{\mathbf{x}}_{t-1}^{(m)}, \tilde{w}_{t-1}^{(m)}\}_{m=1}^{M} \Rightarrow \hat{p}^M(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1}) = \sum_{m=1}^{M} \tilde{w}_{t-1}^{(m)} \delta(\mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{(m)})$
- At each time step t and for $m = 1, ..., M$
  1. Propagate (**Prediction**): $\mathbf{x}_t^{(m)} \sim p(\mathbf{x}_t|\tilde{\mathbf{x}}_{t-1}^{(m)})$
  2. Weights calculation (**Update**): $w_t^{(m)} = \tilde{w}_{t-1}^{(m)} p(y_t|\mathbf{x}_t^{(m)})$

     and normalization: $\bar{w}_t^{(m)} = \frac{w_t^{(m)}}{\sum_{m=1}^{M} w_t^{(m)}}$
  3. **Resampling** (optional but necessary):
     - Sample M times from the random measure $\mathcal{X}_t = \{\mathbf{x}_t^{(m)}, \bar{w}_t^{(m)}\}_{m=1}^{M} \Rightarrow \hat{p}^M(\mathbf{x}_t|\mathbf{y}_{1:t}) = \sum_{m=1}^{M} \bar{w}_t^{(m)} \delta(\mathbf{x}_t - \mathbf{x}_t^{(m)})$
     - New random measure $\tilde{\mathcal{X}}_t$ of M particles with equal weights.

Bootstrap Particle Filter


$\hat{p}^M(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$
3. Resampling
x5
x1
x2
x1
x3
$t-1$
1. Propagate (predict)
$x_t^{(m)} \sim p(x_t|\tilde{x}_{t-1}^{(m)})$
2. Weights calc. (update)
t
$w_t^{(m)} = p(y_t|x_t^{(m)})$
$\hat{p}^M(\mathbf{x}_t|\mathbf{y}_{1:t})$
3. Resampling

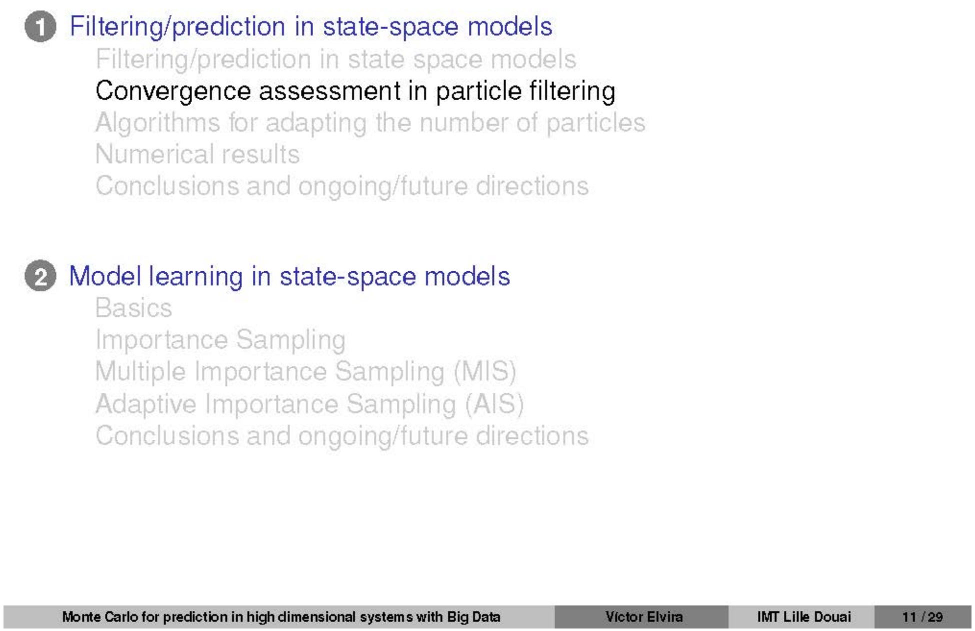
Quality of the Approximation

- The integral of interest is **approximated** $(f, p_t) \approx (f, p_t^M)$ as

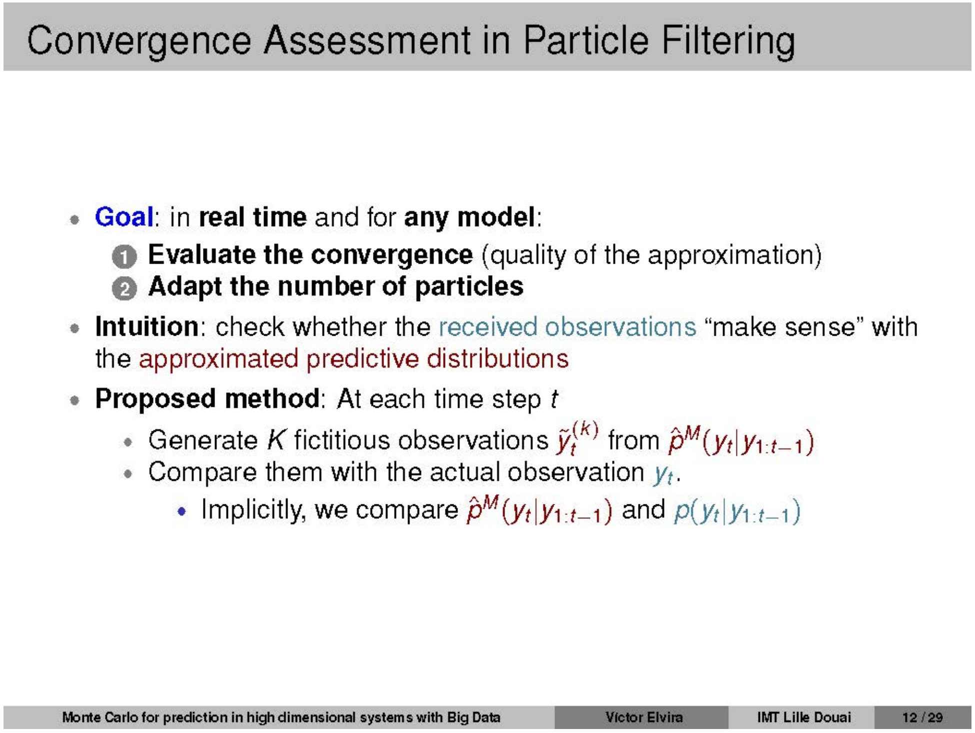
$$\int f(\mathbf{x}_t)p(\mathbf{x}_t|y_{1:t})d\mathbf{x}_t \approx \sum_{m=1}^{M} \bar{w}_t^{(m)} f(\mathbf{x}_t^{(m)})$$

- Some **convergence results** under regularity assumptions:
  - Limit: $\lim_{M\to\infty} |(f, p_t^M - (f, p_t)| = 0$ a.s.
  - Convergence rate: $\mathbb{E}\left[\left((f, p_t^M) - (f, p_t)\right)^2\right] \leq \frac{c_t ||f||_\infty}{M}$
- Peformance/computational cost **tradeoff**
  - Very large M: good approximation but very expensive
  - Reducing M: deteriorates the performance

So, how is M chosen?

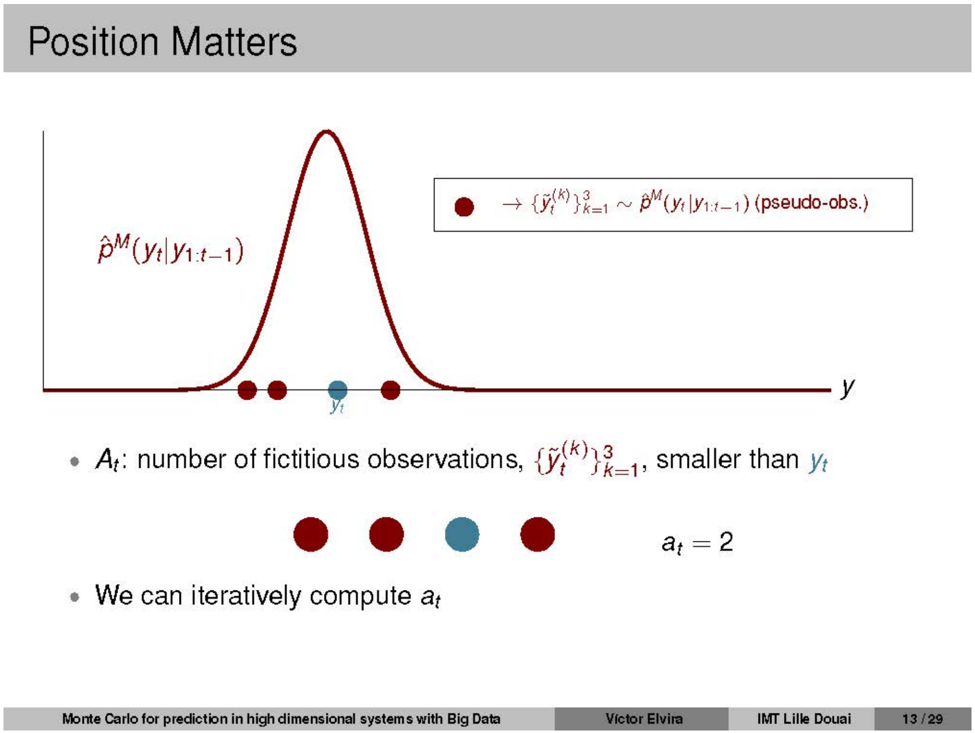


1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions
2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



Convergence Assessment in Particle Filtering

- **Goal**: in **real time** and for **any model**:
  1. **Evaluate the convergence** (quality of the approximation)
  2. **Adapt the number of particles**
- **Intuition**: check whether the received observations “make sense” with the approximated predictive distributions
- **Proposed method**: At each time step t
  - Generate K fictitious observations $\tilde{y}_t^{(k)}$ from $\hat{p}^M(y_t|y_{1:t-1})$
  - Compare them with the actual observation y_t.
    - Implicitly, we compare $\hat{p}^M(y_t|y_{1:t-1})$ and $p(y_t|y_{1:t-1})$

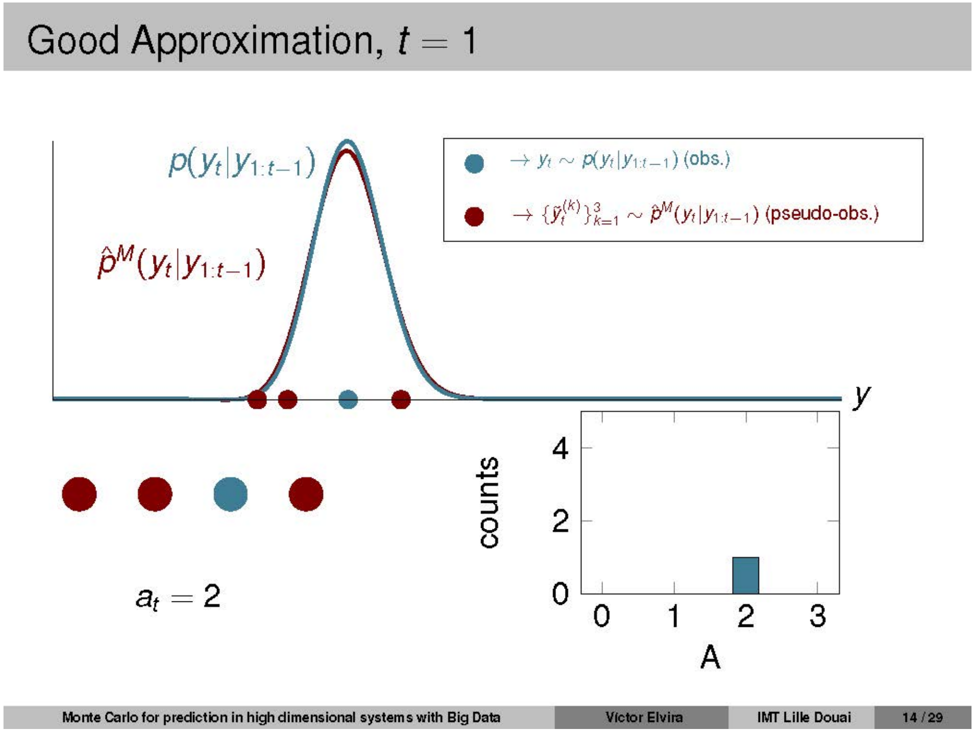


Position Matters

- A_t: number of fictitious observations, $\{\tilde{y}_t^{(k)}\}_{k=1}^3$, smaller than y_t

● ● ● ● $a_t = 2$

- We can iteratively compute a_t

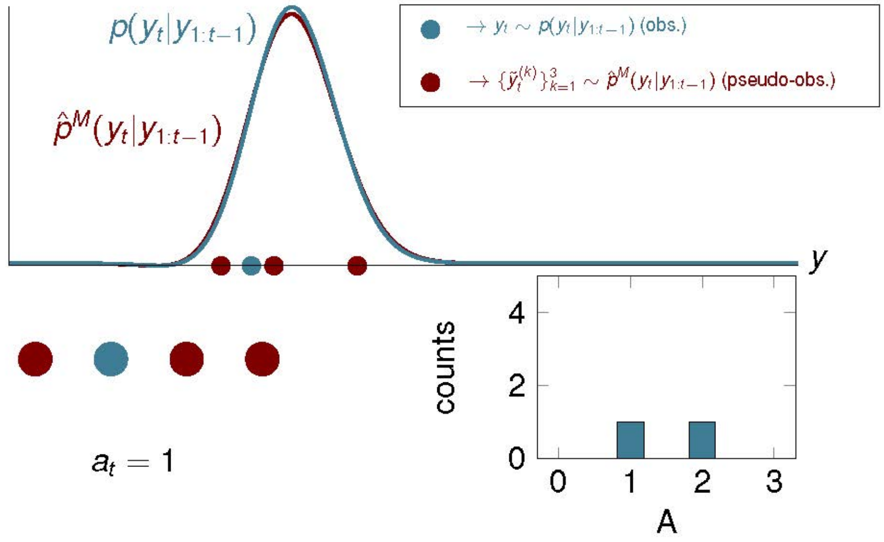


Good Approximation, $t = 1$

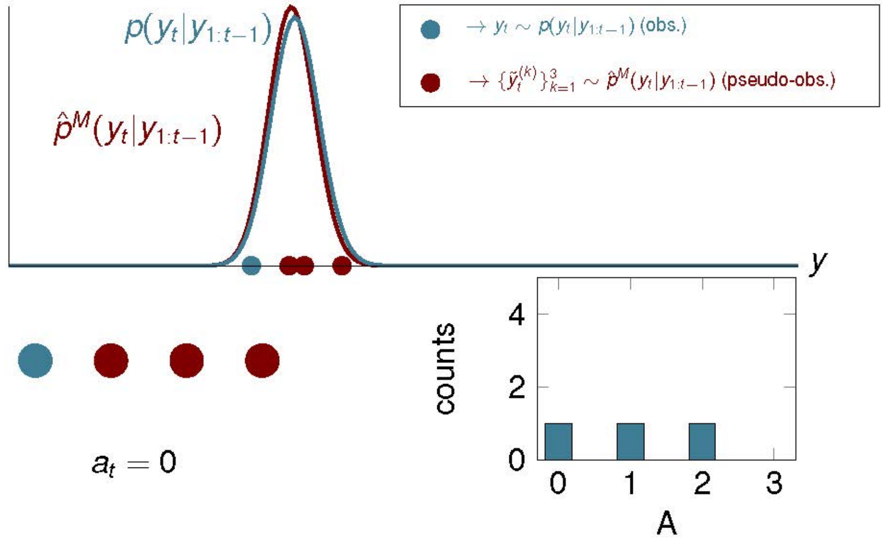
● ● ● ●

$a_t = 2$

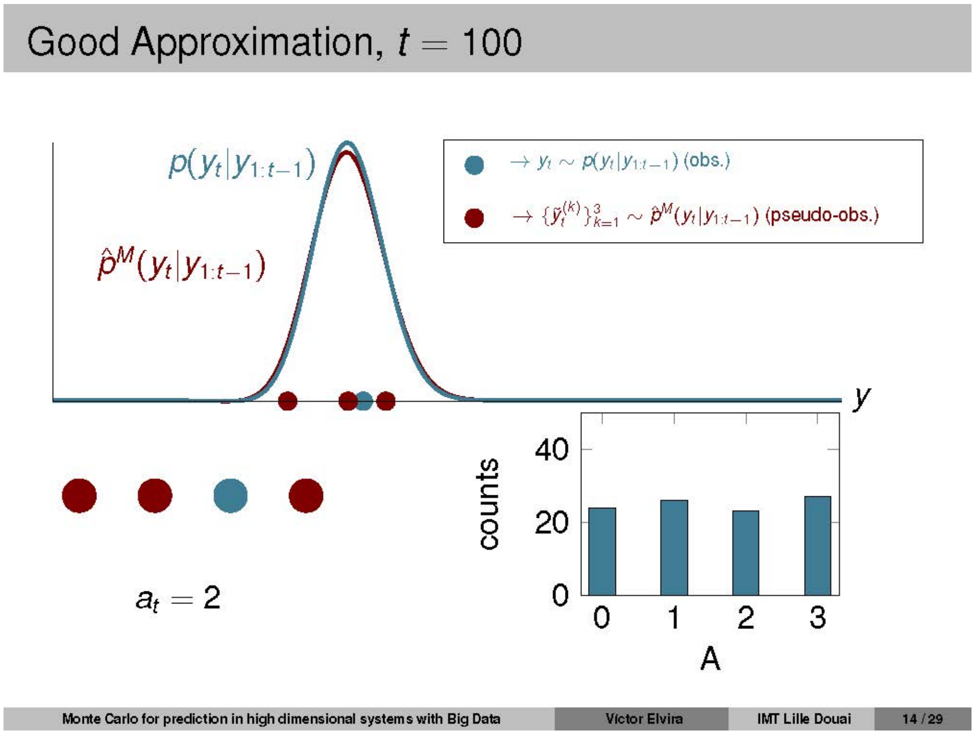
Good Approximation, $t = 2$



Good Approximation, $t = 3$

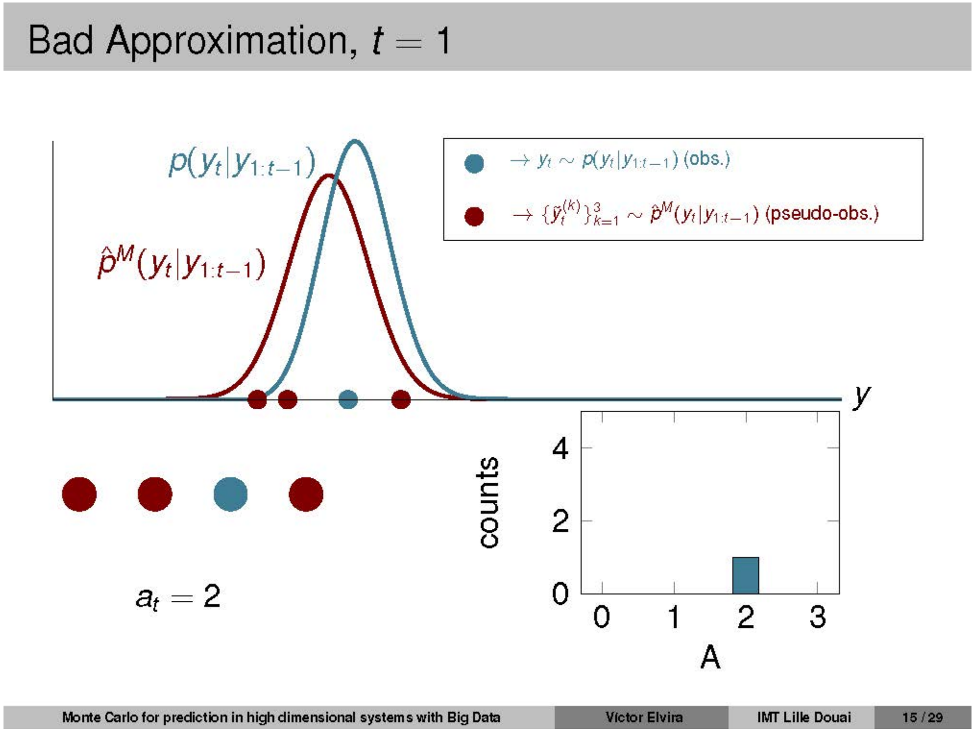


Good Approximation, $t = 100$

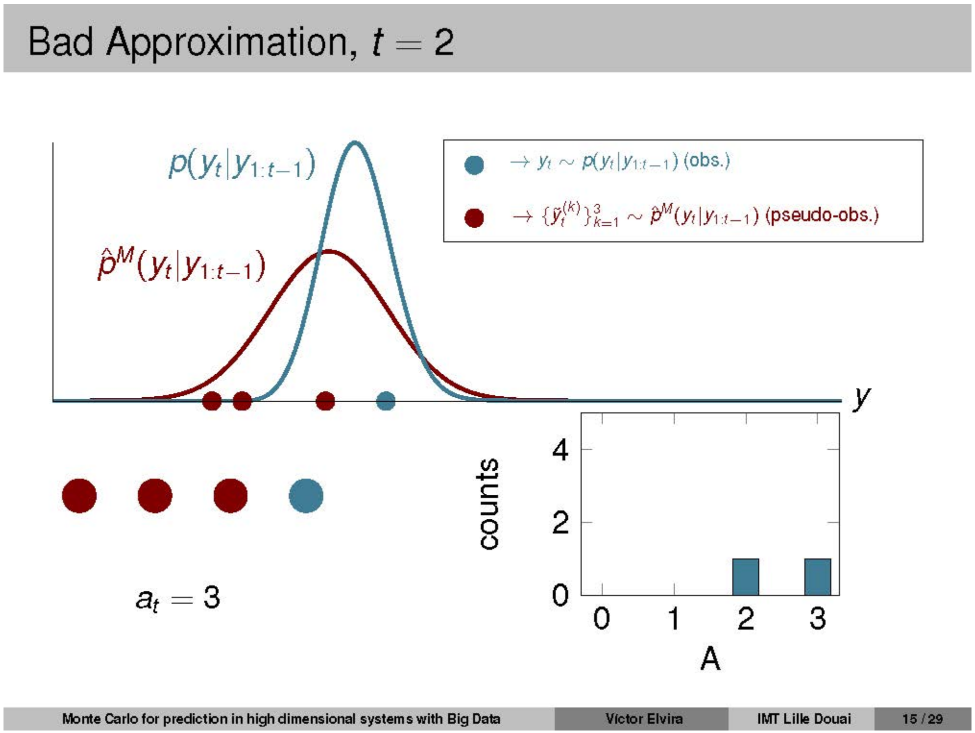


Monte Carlo for prediction in high dimensional systems with Big Data | Víctor Elvira | IMT Lille Douai | 14 / 29

Bad Approximation, $t = 1$



Monte Carlo for prediction in high dimensional systems with Big Data | Víctor Elvira | IMT Lille Douai | 15 / 29

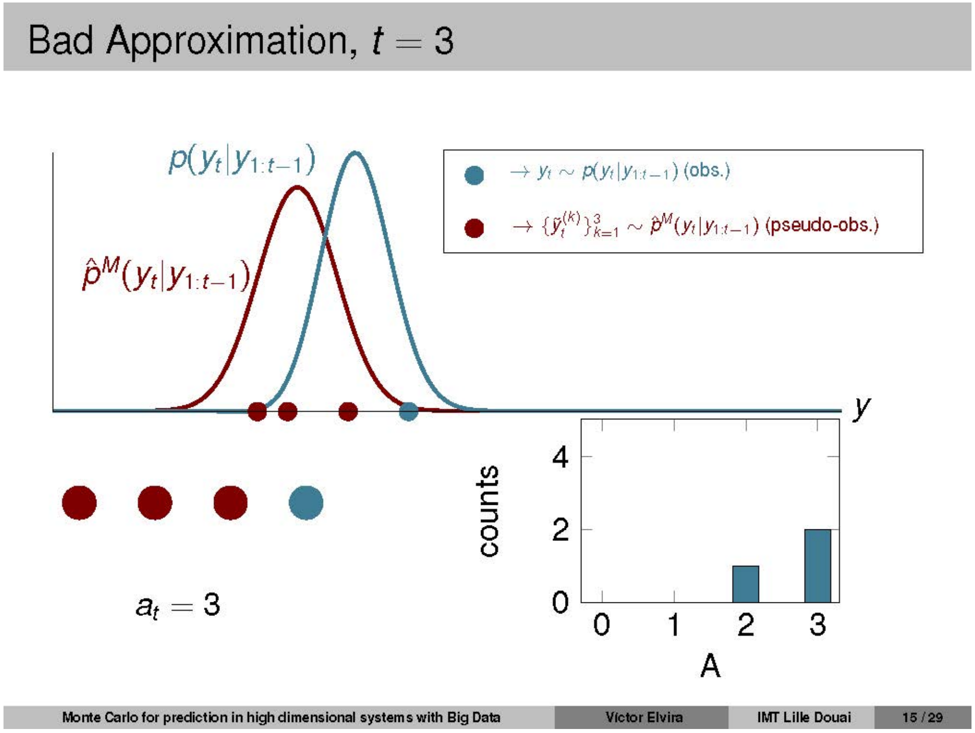


Bad Approximation, $t = 2$

● $\rightarrow y_t \sim p(y_t|y_{1:t-1})$ (obs.)

● $\rightarrow \{\tilde{y}_t^{(k)}\}_{k=1}^{3} \sim \hat{p}^M(y_t|y_{1:t-1})$ (pseudo-obs.)

Monte Carlo for prediction in high dimensional systems with Big Data | Víctor Elvira | IMT Lille Douai | 15 / 29



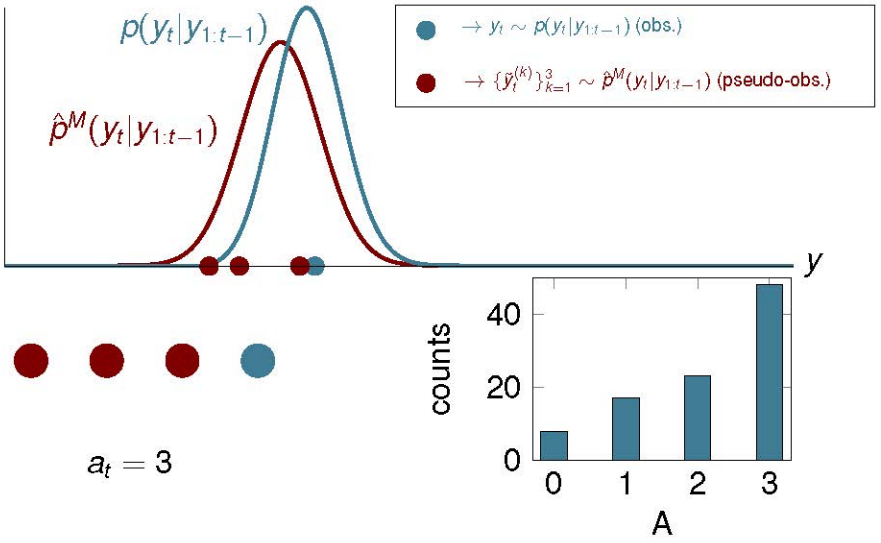
Bad Approximation, $t = 3$

● $\rightarrow y_t \sim p(y_t|y_{1:t-1})$ (obs.)

● $\rightarrow \{\tilde{y}_t^{(k)}\}_{k=1}^{3} \sim \hat{p}^M(y_t|y_{1:t-1})$ (pseudo-obs.)

Monte Carlo for prediction in high dimensional systems with Big Data | Víctor Elvira | IMT Lille Douai | 15 / 29

Bad Approximation, $t = 100$



Methodology Summary and Properties

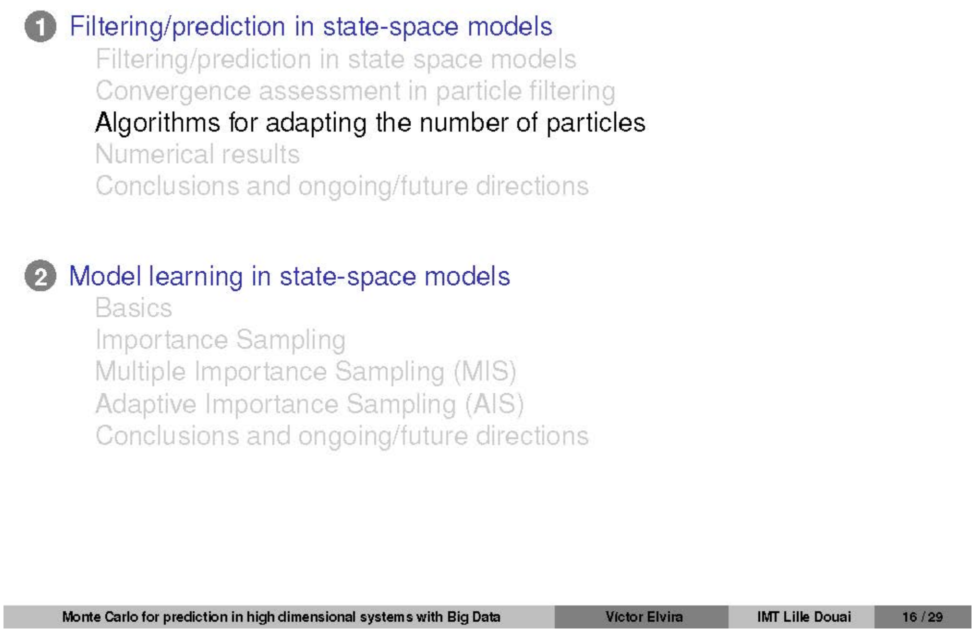
- **Properties**: Under the hypothesis of **perfect approximation**:
  - $\mathcal{J}_t := \{y_t, \tilde{y}_t^{(1)}, \ldots, \tilde{y}_t^{(K)}\}$ is a set of i.i.d. samples from a **common** continuous probability distribution $p_t(y_t)$, then:

Proposition 1: *the pmf of the r.v. $A_{K,t}$ is* ***uniform****:*

$$\mathbb{Q}_K(n) = \frac{1}{K+1}, \qquad n = 0, ..., K.$$

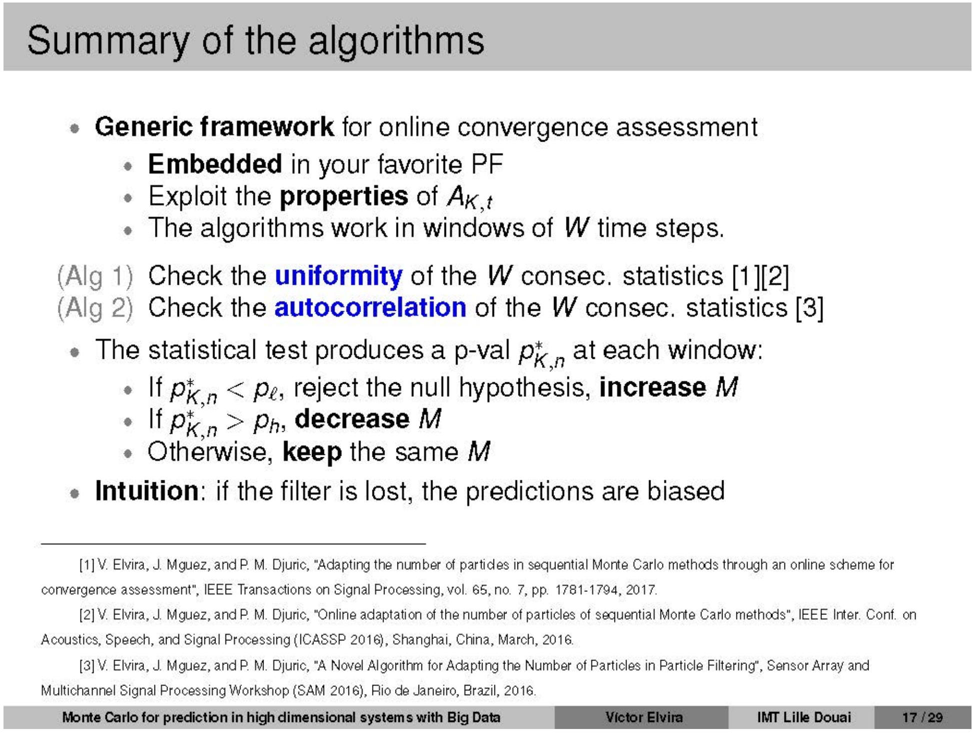
Proposition 2: *the r.v.'s A_{K,t_1} and A_{K,t_2} are* ***independent****, $\forall t_1, t_2 \in \mathbb{N}$ with $t_1 \neq t_2$.*

- **Invariant wrt the state space model!**



1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



Summary of the algorithms

- **Generic framework** for online convergence assessment
  - **Embedded** in your favorite PF
  - Exploit the **properties** of $A_{K,t}$
  - The algorithms work in windows of W time steps.

(Alg 1) Check the **uniformity** of the W consec. statistics [1][2]
(Alg 2) Check the **autocorrelation** of the W consec. statistics [3]

- The statistical test produces a p-val $p^*_{K,n}$ at each window:
  - If $p^*_{K,n} < p_\ell$, reject the null hypothesis, **increase** M
  - If $p^*_{K,n} > p_h$, **decrease** M
  - Otherwise, **keep** the same M
- **Intuition**: if the filter is lost, the predictions are biased

[1] V. Elvira, J. Mguez, and P. M. Djuric, "Adapting the number of particles in sequential Monte Carlo methods through an online scheme for convergence assessment", IEEE Transactions on Signal Processing, vol. 65, no. 7, pp. 1781-1794, 2017.

[2] V. Elvira, J. Mguez, and P. M. Djuric, "Online adaptation of the number of particles of sequential Monte Carlo methods", IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai, China, March, 2016.

[3] V. Elvira, J. Mguez, and P. M. Djuric, "A Novel Algorithm for Adapting the Number of Particles in Particle Filtering", Sensor Array and Multichannel Signal Processing Workshop (SAM 2016), Rio de Janeiro, Brazil, 2016.

1. Filtering/prediction in state-space models
    - Filtering/prediction in state space models
    - Convergence assessment in particle filtering
    - Algorithms for adapting the number of particles
    - Numerical results
    - Conclusions and ongoing/future directions

2. Model learning in state-space models
    - Basics
    - Importance Sampling
    - Multiple Importance Sampling (MIS)
    - Adaptive Importance Sampling (AIS)
    - Conclusions and ongoing/future directions

Alg. 2 (autocorrelation): Stochastic Growth Model

- Stochastic growth model:

$$x_t = \frac{x_{t-1}}{2} + \frac{25x_{t-1}}{1+x_{t-1}^2} + 8\cos(\phi t) + u_t, \quad (1)$$

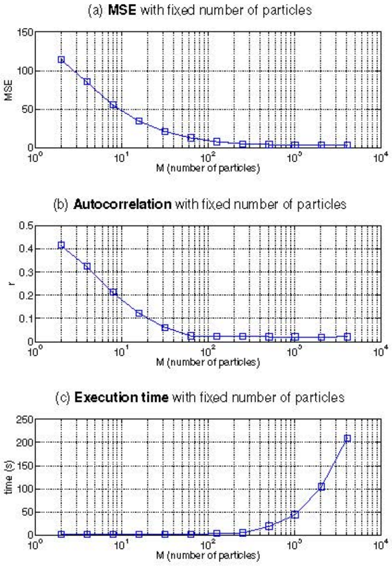
$$y_t = \frac{x_t^2}{20} + v_t, \quad (2)$$

with $\phi = 0.4$, u_t and v_t are i.i.d. zero-mean univariate Gaussian r.v.'s with variance $\sigma_u^2 = 2$ and $\sigma_v^2 = 0.1$. $T = 10^4$.

- Algorithm parameters: $K = 7$, $W = 25$

$[p_l - p_h]$	$[0.2 - 0.6]$	$[0.25 - 0.65]$	$[0.35 - 0.75]$	$[0.4 - 0.8]$	$[0.45 - 0.85]$
MSE	21.62	13.83	4.90	3.62	3.39
M	144	386	1933	2841	3255
ex. time (s)	18.9	233.4	285.7	441.5	536.1

Alg. 2 (autocorrelation): Stochastic Growth Model

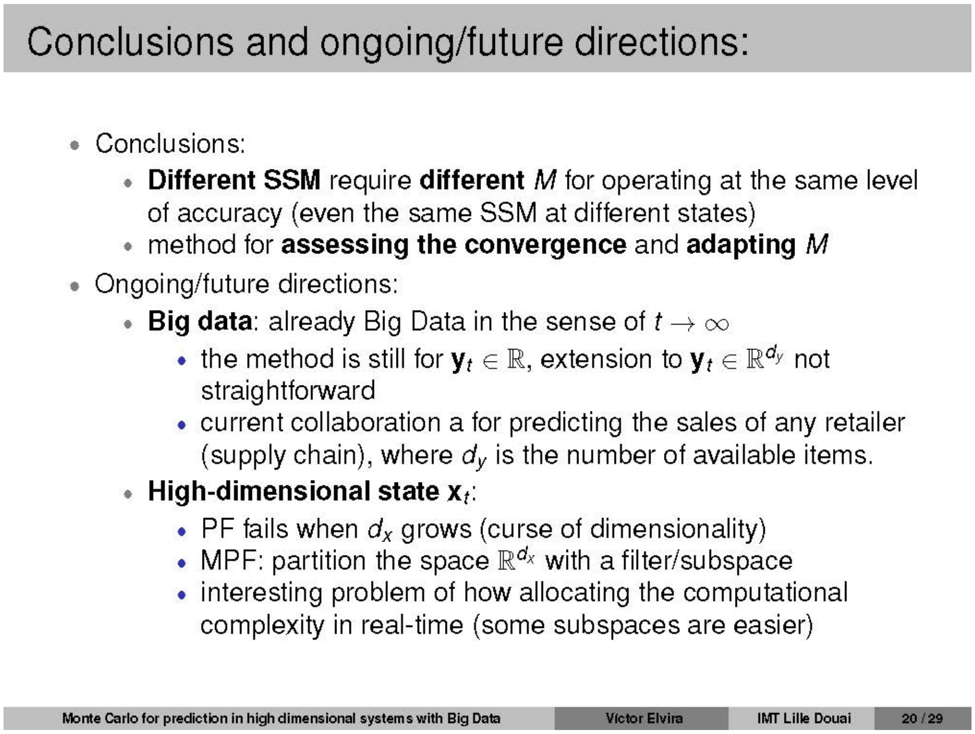

(a) **MSE** with fixed number of particles

(b) **Autocorrelation** with fixed number of particles

(c) **Execution time** with fixed number of particles

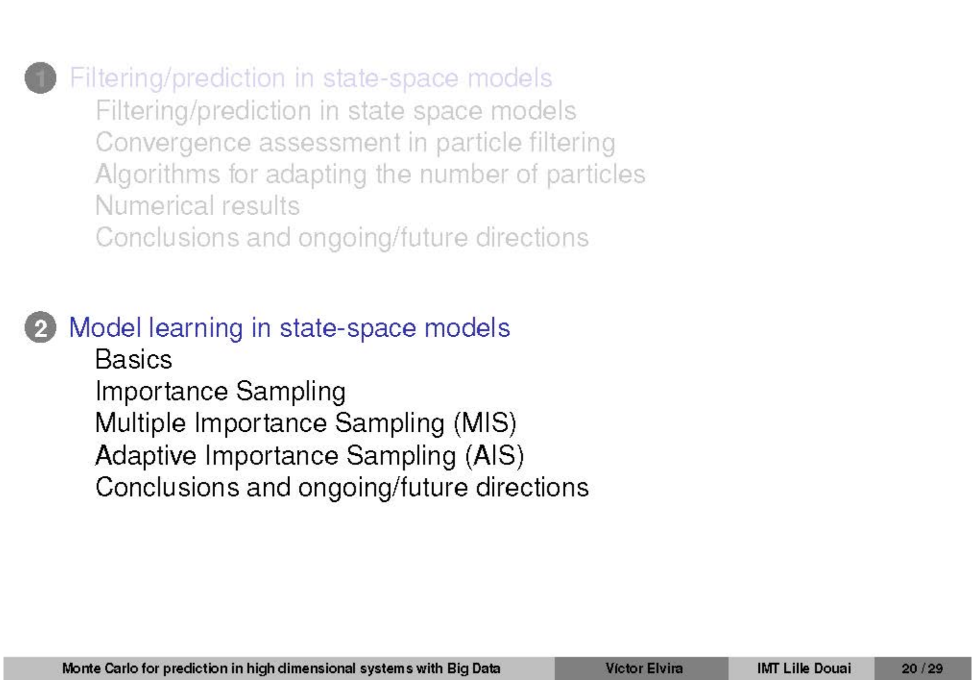
1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



Conclusions and ongoing/future directions:

- Conclusions:
  - **Different SSM** require **different** M for operating at the same level of accuracy (even the same SSM at different states)
  - method for **assessing the convergence** and **adapting** M
- Ongoing/future directions:
  - **Big data**: already Big Data in the sense of $t \to \infty$
    - the method is still for $\mathbf{y}_t \in \mathbb{R}$, extension to $\mathbf{y}_t \in \mathbb{R}^{d_y}$ not straightforward
    - current collaboration a for predicting the sales of any retailer (supply chain), where d_y is the number of available items.
  - **High-dimensional state** $\mathbf{x}_t$:
    - PF fails when d_x grows (curse of dimensionality)
    - MPF: partition the space $\mathbb{R}^{d_x}$ with a filter/subspace
    - interesting problem of how allocating the computational complexity in real-time (some subspaces are easier)



1 Filtering/prediction in state-space models
Filtering/prediction in state space models
Convergence assessment in particle filtering
Algorithms for adapting the number of particles
Numerical results
Conclusions and ongoing/future directions

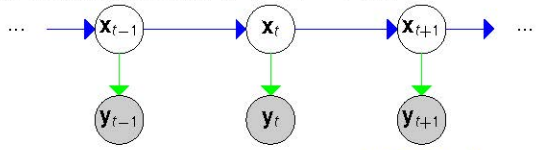
2 Model learning in state-space models
Basics
Importance Sampling
Multiple Importance Sampling (MIS)
Adaptive Importance Sampling (AIS)
Conclusions and ongoing/future directions

1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions

The inference/prediction problem under unknown model

- Let us consider:
  - a set of hidden states $\mathbf{x}_t \in \mathbb{R}^{d_x}$, $t = 1, ..., T$.
  - a set of observations $\mathbf{y}_t \in \mathbb{R}^{d_y}$, $t = 1, ..., T$.



$$\begin{aligned}
\mathbf{x}_t &= g_\theta(\mathbf{x}_{t-1}, \mathbf{u}_t) \rightarrow && p_\theta(\mathbf{x}_t|\mathbf{x}_{t-1}) \\
\mathbf{y}_t &= h_\theta(\mathbf{x}_t, \mathbf{v}_t) \rightarrow && p_\theta(\mathbf{y}_t|\mathbf{x}_t)
\end{aligned}$$

g_θ and h_θ are known and $\mathbf{u}_t$ and $\mathbf{v}_t$ have known distribution, but

θ is unknown $\equiv$ realistic scenario

1. Batch approach
2. Sequential approach (ongoing/future)

Problem Statement: Batch Approach

- All observations up to $t = T$ are processed together $\mathbf{y}_{1:T}$
- **Probabilistic inference over the θ as well**. Bayes rule:

$$p(\boldsymbol{\theta}|\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_{1:T}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y}_{1:T})} \equiv \frac{\pi(\boldsymbol{\theta})}{p(\mathbf{y}_{1:T})}$$

  - $p(\boldsymbol{\theta}|\mathbf{y}_{1:T}) \equiv \tilde{\pi}(\boldsymbol{\theta})$ is our target
  - $p(\mathbf{y}_{1:T}|\boldsymbol{\theta}) = \int p(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}|\boldsymbol{\theta})d\mathbf{x}_{1:T}$ is not available in close-form and must be approximated
  - $p(\mathbf{y}_{1:T})$ must be approximated
- We want to approximate integrals of the form

$$I = \int g(\boldsymbol{\theta})\tilde{\pi}(\boldsymbol{\theta})d\boldsymbol{\theta}$$

e.g., $g(\boldsymbol{\theta}) = \boldsymbol{\theta}$ for the mean.

1 Filtering/prediction in state-space models
   Filtering/prediction in state space models
   Convergence assessment in particle filtering
   Algorithms for adapting the number of particles
   Numerical results
   Conclusions and ongoing/future directions

2 Model learning in state-space models
   Basics
   Importance Sampling
   Multiple Importance Sampling (MIS)
   Adaptive Importance Sampling (AIS)
   Conclusions and ongoing/future directions

Importance Sampling: Basics

- Limitations in our problem:
  - We cannot evaluate nor sample from $\tilde{\pi}(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$
  - We cannot evaluate $\pi(\boldsymbol{\theta}) = p(\mathbf{y}_{1:T}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
    - we can approximately evaluate $\hat{p}(\mathbf{y}_{1:T}|\boldsymbol{\theta})$ in a point $\boldsymbol{\theta}_n$ running the PF with $\boldsymbol{\theta} = \boldsymbol{\theta}_n$, $\pi(\boldsymbol{\theta}_n) \approx \hat{\pi}(\boldsymbol{\theta}_n) = \hat{p}(\mathbf{y}_{1:T}|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n)$
- Two basic steps in IS:
  1. **Sampling:** N samples from the proposal $q(\boldsymbol{\theta})$ instead of $\tilde{\pi}(\boldsymbol{\theta})$

$$\boldsymbol{\theta}_n \sim q(\boldsymbol{\theta}), \qquad n = 1, ..., N.$$

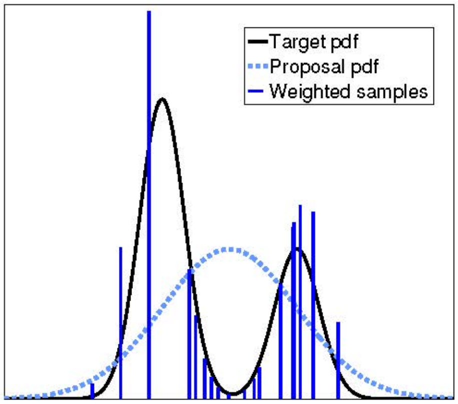
  2. **Weighting:** Assigned with an importance weight

$$w_n = \frac{\hat{\pi}(\boldsymbol{\theta}_n)}{q(\boldsymbol{\theta}_n)}, \qquad n = 1, ..., N.$$

$$I = \int g(\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{y}_{1:T})d\boldsymbol{\theta} \approx \widetilde{I} = \frac{1}{N}\sum_{n=1}^{N} \bar{w}_n g(\boldsymbol{\theta}_n), \quad \bar{w}_n = \frac{w_n}{\sum_{i=1}^{N} w_i}$$

Remark: Each evaluation $\hat{\pi}(\boldsymbol{\theta}_n)$ requires re-processing all observations $\mathbf{y}_{1:t}$.

Importance Sampling: Basics (II)

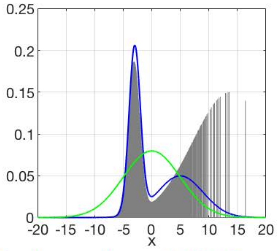


- Target pdf: $p(\boldsymbol{\theta}|\mathbf{y}_{1:T}) \approx \frac{1}{N}\sum_{n=1}^{N} \bar{w}_n \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_n)$
- Proposal pdf: $q(\boldsymbol{\theta})$
- Weighted samples: $\{\boldsymbol{\theta}_n, \bar{w}_n\}_{n=1}^{N}$

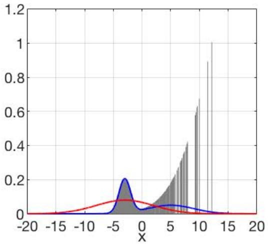
Importance Sampling: example

- Variance increases with the mismatch of $|g(\boldsymbol{\theta})|\tilde{\pi}(\boldsymbol{\theta})$ and $q(\boldsymbol{\theta})$.

$$\text{Var}_{q(\boldsymbol{\theta})}(\widehat{I}) = \frac{1}{N}\int \frac{g^2(\boldsymbol{\theta})\tilde{\pi}^2(\boldsymbol{\theta})}{q(\boldsymbol{\theta})} d\boldsymbol{\theta} - \frac{I}{N}$$



Good proposal $q_1 = \mathcal{N}(0, 25)$



Bad proposal $q_2 = \mathcal{N}(-3, 25)$

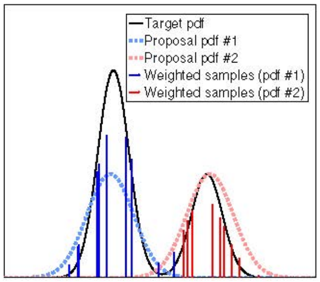
- The approximation converges in both cases, but a good proposal is key for the efficiency of IS.
- Usually difficult to find a good $q(\boldsymbol{\theta})$ because $\tilde{\pi}(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$:
  - evaluated only up to a normalizing constant $\pi(\boldsymbol{\theta}) = p(\mathbf{y}_{1:T}|\boldsymbol{\theta})p(\boldsymbol{\theta})$
  - even the evaluation is approximated as $\hat{\pi}(\boldsymbol{\theta})$ by the PF
- **Exploring** (**adaptivitity**) and **harvesting** with **multiple** proposals.

1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions

Multiple Importance Sampling (MIS): Basics

- Set of N proposal pdfs $\{q_1(\theta), q_2(\theta), ..., q_N(\theta)\}$ available.

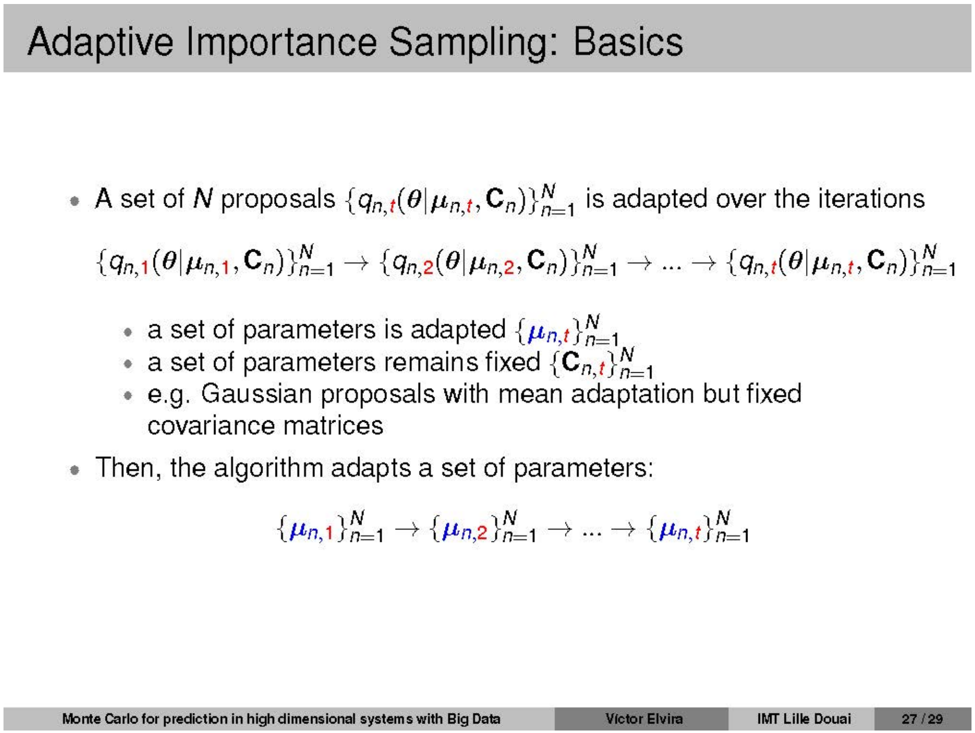


- Extension from unique proposal to MIS is **not trivial**
- In [Elvira16]: Novel framework for valid **sampling**/**weighting** in MIS
  - 3 specific sampling procedures
  - 5 generic weigthing functions (not add-hoc, but related to distributions of the sampling procedure)
  - Variance comparison among the 6 MIS schemes

[Elvira16] V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Generalized Multiple Importance Sampling", arXiv:1511.03095, 2016.

1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



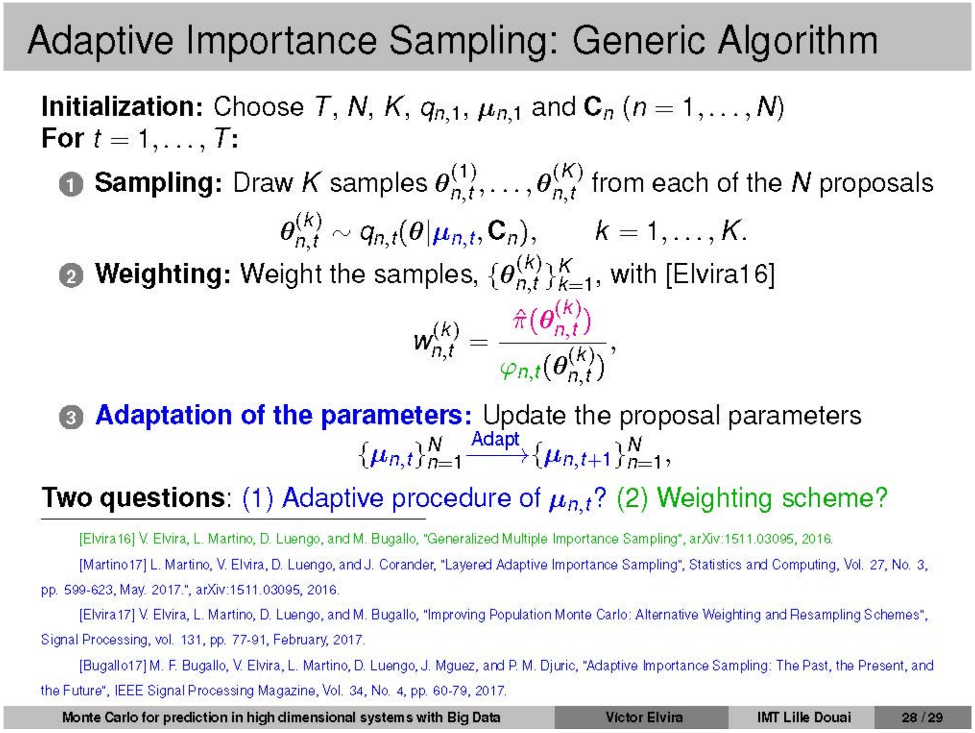
Adaptive Importance Sampling: Basics

- A set of N proposals $\{q_{n,t}(\theta|\mu_{n,t},\mathbf{C}_n)\}_{n=1}^N$ is adapted over the iterations

$$\{q_{n,1}(\theta|\mu_{n,1},\mathbf{C}_n)\}_{n=1}^N \to \{q_{n,2}(\theta|\mu_{n,2},\mathbf{C}_n)\}_{n=1}^N \to ... \to \{q_{n,t}(\theta|\mu_{n,t},\mathbf{C}_n)\}_{n=1}^N$$

  - a set of parameters is adapted $\{\mu_{n,t}\}_{n=1}^N$
  - a set of parameters remains fixed $\{\mathbf{C}_{n,t}\}_{n=1}^N$
  - e.g. Gaussian proposals with mean adaptation but fixed covariance matrices
- Then, the algorithm adapts a set of parameters:

$$\{\mu_{n,1}\}_{n=1}^N \to \{\mu_{n,2}\}_{n=1}^N \to ... \to \{\mu_{n,t}\}_{n=1}^N$$



Adaptive Importance Sampling: Generic Algorithm

Initialization: Choose T, N, K, $q_{n,1}$, $\mu_{n,1}$ and $\mathbf{C}_n$ ($n = 1, \dots, N$)
For $t = 1, \dots, T$**:**

1. **Sampling:** Draw K samples $\theta_{n,t}^{(1)}, \dots, \theta_{n,t}^{(K)}$ from each of the N proposals
$$\theta_{n,t}^{(k)} \sim q_{n,t}(\theta|\mu_{n,t},\mathbf{C}_n), \qquad k = 1, \dots, K.$$
2. **Weighting:** Weight the samples, $\{\theta_{n,t}^{(k)}\}_{k=1}^K$, with [Elvira16]
$$w_{n,t}^{(k)} = \frac{\hat{\pi}(\theta_{n,t}^{(k)})}{\varphi_{n,t}(\theta_{n,t}^{(k)})},$$
3. **Adaptation of the parameters:** Update the proposal parameters
$$\{\mu_{n,t}\}_{n=1}^N \xrightarrow{\text{Adapt}} \{\mu_{n,t+1}\}_{n=1}^N,$$

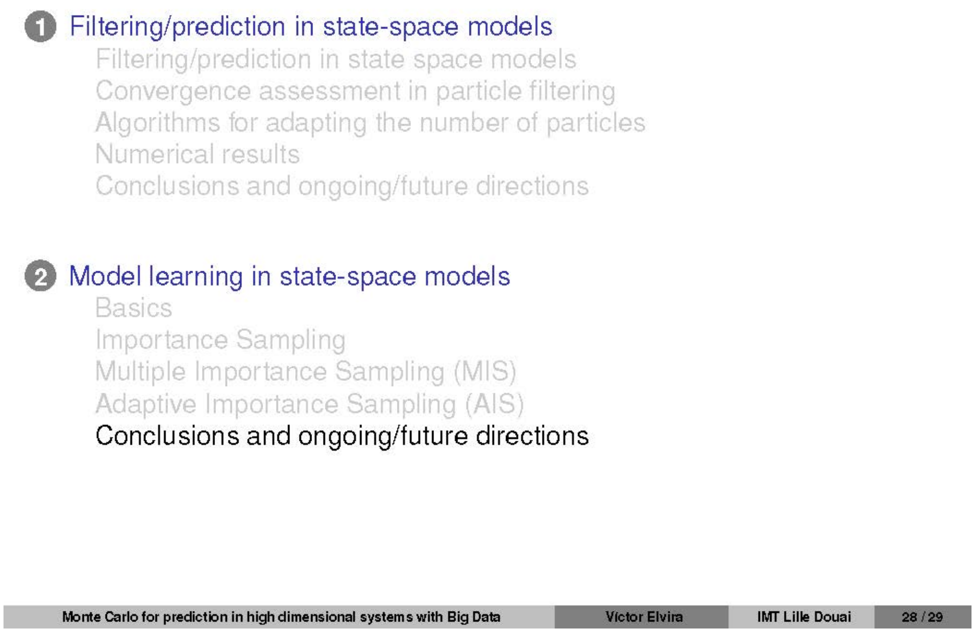
Two questions: (1) Adaptive procedure of $\mu_{n,t}$? (2) Weighting scheme?

[Elvira16] V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Generalized Multiple Importance Sampling", arXiv:1511.03095, 2016.

[Martino17] L. Martino, V. Elvira, D. Luengo, and J. Corander, "Layered Adaptive Importance Sampling", Statistics and Computing, Vol. 27, No. 3, pp. 599-623, May. 2017.", arXiv:1511.03095, 2016.

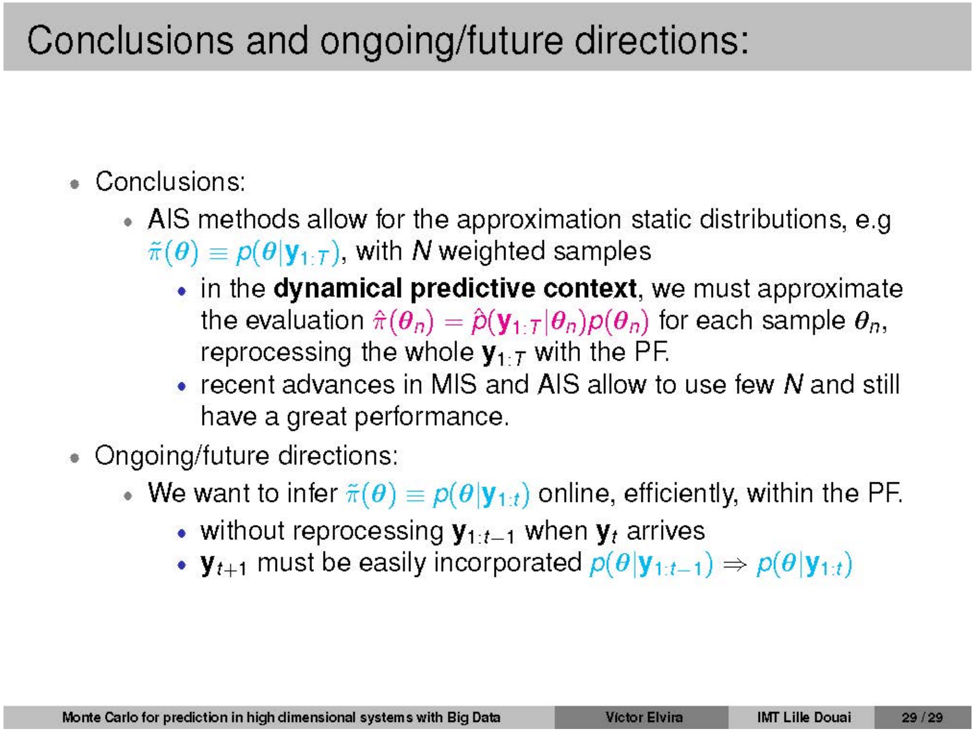
[Elvira17] V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Improving Population Monte Carlo: Alternative Weighting and Resampling Schemes", Signal Processing, vol. 131, pp. 77-91, February, 2017.

[Bugallo17] M. F. Bugallo, V. Elvira, L. Martino, D. Luengo, J. Mguez, and P. M. Djuric, "Adaptive Importance Sampling: The Past, the Present, and the Future", IEEE Signal Processing Magazine, Vol. 34, No. 4, pp. 60-79, 2017.



1. Filtering/prediction in state-space models
   - Filtering/prediction in state space models
   - Convergence assessment in particle filtering
   - Algorithms for adapting the number of particles
   - Numerical results
   - Conclusions and ongoing/future directions

2. Model learning in state-space models
   - Basics
   - Importance Sampling
   - Multiple Importance Sampling (MIS)
   - Adaptive Importance Sampling (AIS)
   - Conclusions and ongoing/future directions



Conclusions and ongoing/future directions:

- Conclusions:
  - AIS methods allow for the approximation static distributions, e.g $\tilde{\pi}(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$, with N weighted samples
    - in the **dynamical predictive context**, we must approximate the evaluation $\hat{\tilde{\pi}}(\boldsymbol{\theta}_n) = \hat{p}(\mathbf{y}_{1:T}|\boldsymbol{\theta}_n)p(\boldsymbol{\theta}_n)$ for each sample $\boldsymbol{\theta}_n$, reprocessing the whole $\mathbf{y}_{1:T}$ with the PF.
    - recent advances in MIS and AIS allow to use few N and still have a great performance.
- Ongoing/future directions:
  - We want to infer $\tilde{\pi}(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta}|\mathbf{y}_{1:t})$ online, efficiently, within the PF.
    - without reprocessing $\mathbf{y}_{1:t-1}$ when $\mathbf{y}_t$ arrives
    - $\mathbf{y}_{t+1}$ must be easily incorporated $p(\boldsymbol{\theta}|\mathbf{y}_{1:t-1}) \Rightarrow p(\boldsymbol{\theta}|\mathbf{y}_{1:t})$

Thank you for your attention!

References

- M. F. Bugallo, V. Elvira, L. Martino, D. Luengo, J. Mguez, and P. M. Djuric, "Adaptive Importance Sampling: The Past, the Present, and the Future", IEEE Signal Processing Magazine, Vol. 34, No. 4, pp. 60-79, 2017.
- O. Cappé, A. Guillin, J. M. Marin, and C. P. Robert, "Population Monte Carlo", Journal of Computational and Graphical Statistics, pp. 907-929, 2004.
- O. Cappé, R. Douc, A. Guillin, J. M. Marin, and C. P. Robert, "Adaptive importance sampling in general mixture classes", Statistics and Computing, vol. 18-4, pp. 447-459, 2008.
- J. Cornuet, J. M. Marin, A. Mira, and C. P. Robert, "Adaptive multiple importance sampling. Scandinavian Journal of Statistics", vol. 39-4, pp. 798-812, 2012.
- R. Douc, A. Guillin, J. M. Marin, and C. P. Robert, "Minimum variance importance sampling via population Monte Carlo", ESAIM: Probability and Statistics, vol. 11, pp. 427-447, 2007.
- R. Douc and E. Moulines, "Limit theorems for weighted samples with applications to sequential Monte Carlo methods", in ESAIM: Proceedings, vol. 19, pp. 101-107, 2007.
- V. Elvira, J. Mguez, and P. M. Djuric, "Adapting the number of particles in sequential Monte Carlo methods through an online scheme for convergence assessment", IEEE Transactions on Signal Processing, vol. 65, no. 7, pp. 1781-1794, 2017.
- V. Elvira, J. Mguez, and P. M. Djuric, "Online adaptation of the number of particles of sequential Monte Carlo methods", IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai, China, March, 2016.
- V. Elvira, J. Mguez, and P. M. Djuric, "A Novel Algorithm for Adapting the Number of Particles in Particle Filtering", Sensor Array and Multichannel Signal Processing Workshop (SAM 2016), Rio de Janeiro, Brazil, 2016.
- V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Efficient Multiple Importance Sampling Estimators", IEEE Signal Processing Letters, vol. 22, no. 10, pp. 1757-1761, March, 2015.
- V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Generalized Multiple Importance Sampling", arXiv:1511.03095, 2016.
- V. Elvira, L. Martino, D. Luengo, and M. F. Bugallo, "Novel weighting schemes with non-disjoint mixtures of proposals in multiple importance sampling", IEEE Workshop on Statistical Signal Processing (SSP 2016), Mallorca, Spain, June, 2016.
- V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Heretical Multiple Importance Sampling", IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1474-1478, October, 2016.
- V. Elvira, L. Martino, D. Luengo, and M. Bugallo, "Improving Population Monte Carlo: Alternative Weighting and Resampling Schemes", Signal Processing, vol. 131, pp. 77-91, February, 2017.
- L. Martino, V. Elvira, D. Luengo, and J. Corander, "An Adaptive Population Importance Sampler: Learning from the Uncertainty", IEEE Transactions on Signal Processing, vol. 63, no. 16, pp. 4422-4437, August, 2015.
- L. Martino, V. Elvira, D. Luengo, and J. Corander, "Layered Adaptive Importance Sampling", Statistics and Computing, Vol. 27, No. 3, pp. 599-623, May. 2017.
- A. Owen and Y. Zhou, "Safe and effective importance sampling", Journal of the American Statistical Association, vol. 95-449, pp. 135-143, 2000.
- C. P. Robert and G. Casella, "Monte Carlo statistical methods", Springer Science and Business Media, 2004.

Monte Carlo for prediction in high dimensional systems with Big Data | Víctor Elvira | IMT Lille Douai | 29 / 29

Theoretical Results

- **Theoretical analysis**:
  - **convergence** of the predictive pdf of the observations:

$$\lim_{M\to\infty} \Big(f, \hat{p}^M(y_t|y_{1:t-1})\Big) = \Big(f, p(y_t|y_{1:t-1})\Big) \qquad \text{a.s.},$$

    with explicit **convergence rate**
    - extends the existing results of pointwise convergence of $\hat{p}^M(y_t|y_{1:t-1})$ to $\hat{p}(y_t|y_{1:t-1})$
    - holds for **multidimensional** observations
    - key for the statistical analysis of $A_{K,t}$
  - **convergence** of the p.m.f. of $A_{K,t}$ to a discrete uniform distribution

$$\frac{1}{K+1} - \varepsilon_M \leq \mathbb{Q}_K(n) \leq \frac{1}{K+1} + \varepsilon_M, \qquad n = 0, ..., K,$$

    with $\lim_{M\to\infty} \varepsilon_M = 0$ a.s.

[1] V. Elvira, J. Mguez, and P. M. Djuric, "Adapting the number of particles in sequential Monte Carlo methods through an online scheme for convergence assessment", IEEE Transactions on Signal Processing, vol. 65, no. 7, pp. 1781-1794, 2017.

Alg. 1 (uniformity): Lorenz63 System

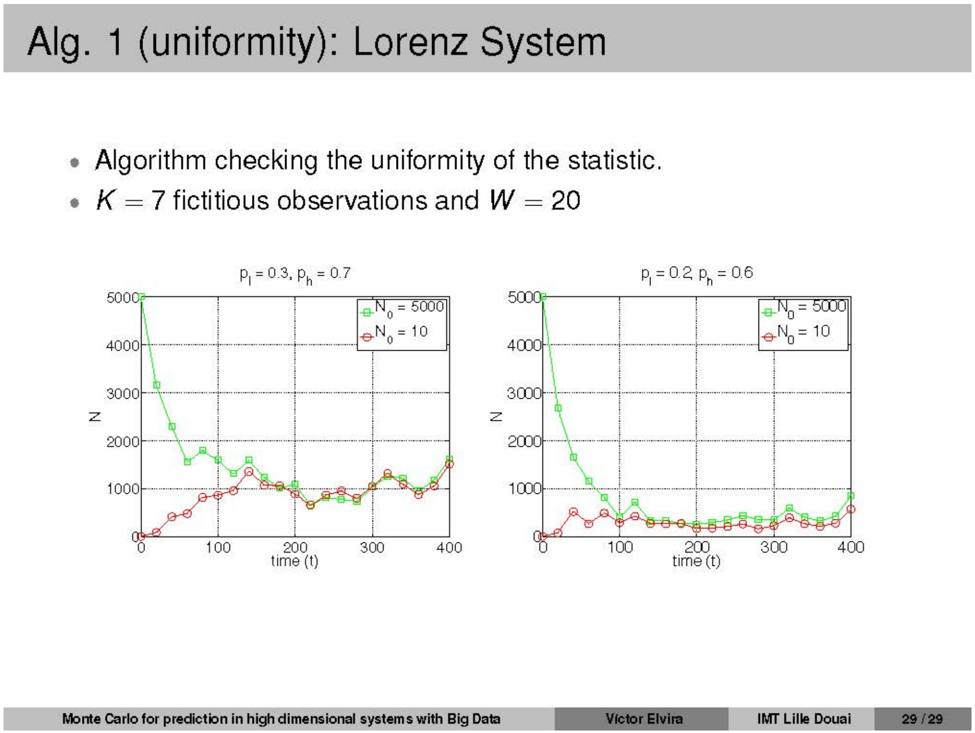
- 3-dimensional dynamical system defined by

$$\begin{aligned} dX_1 &= -\mathrm{s}(X_1 - X_2), \\ dX_2 &= \mathrm{r}X_1 - X_2 - X_1X_3, \\ dX_3 &= X_1X_2 - \mathrm{b}X_3, \end{aligned}$$

- Time discrete version using Euler's method with

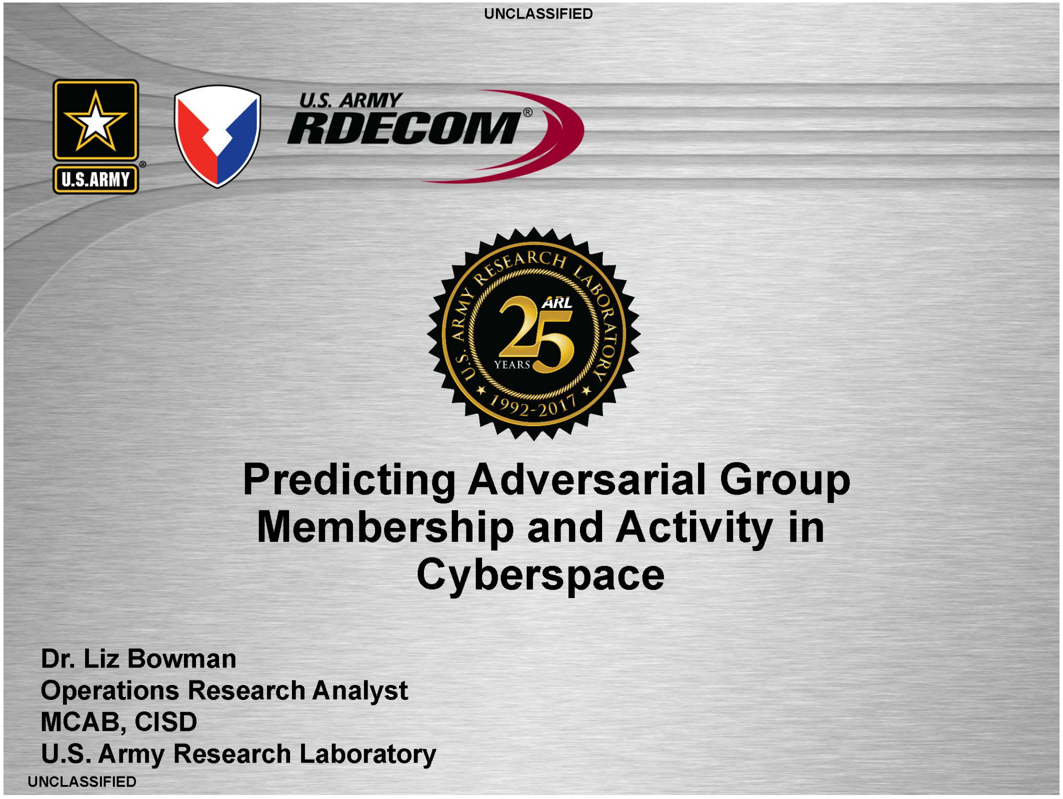
$$\begin{aligned} X_{1,n} &= X_{1,n-1} - \Delta\mathrm{s}(X_{1,n-1} - X_{2,n-1}) + \sqrt{\Delta}U_{1,n}, \\ X_{2,n} &= X_{2,n-1} + \Delta(\mathrm{r}X_{1,n-1} - X_{2,n-1} - X_{1,n-1}X_{3,n-1}) + \sqrt{\Delta}U_{2,n}, \\ X_{3,n} &= X_{3,n-1} + \Delta(X_{1,n-1}X_{2,n-1} - \mathrm{b}X_{3,n-1}) + \sqrt{\Delta}U_{3,n}, \end{aligned}$$

- $U_{i,n} \sim \mathcal{N}(0,1)$, $\Delta = 10^{-3}$, and $(\mathrm{s},\mathrm{r},\mathrm{b}) = \left(10, 28, \frac{8}{3}\right)$



Alg. 1 (uniformity): Lorenz System

- Algorithm checking the uniformity of the statistic.
- $K = 7$ fictitious observations and $W = 20$



UNCLASSIFIED

U.S. ARMY
U.S. ARMY RDECOM®
U.S. ARMY RESEARCH LABORATORY
ARL
25
YEARS
1992-2017

Predicting Adversarial Group Membership and Activity in Cyberspace

Dr. Liz Bowman
Operations Research Analyst
MCAB, CISD
U.S. Army Research Laboratory

UNCLASSIFIED

Presentation Outline

- Increasing complexity of information retrieval in complex and data-dense asymmetric environments
- Social Understanding and Reasoning Framework (SURF) tool development and analysis tasks
  - Analyse the social network,
  - Identify the key influencers, and
  - Analyse the structural elements of message traffic between key influencers
- Scenario
- Case study results
- Conclusions
- Next steps

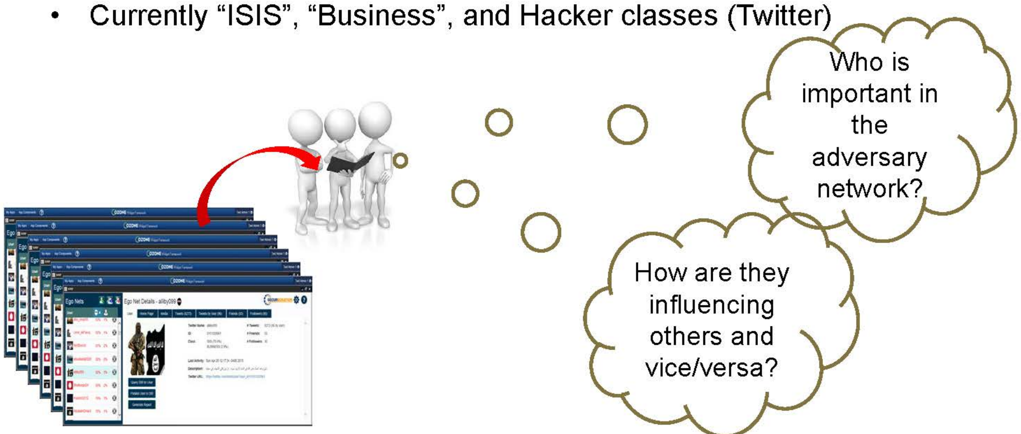
Increasing Complexity of Military Challenges and Expanding Cache of Interesting Data

The Problem:

- Too much data is a common complaint in most operational domains
- This limitation requires decision makers to mentally reconstruct, infer, and extract relevant information through laborious and error-prone internal processes
- Automated mechanisms are needed for the timely extraction & prioritization of high-value decision-relevant information

SURF Tool Development

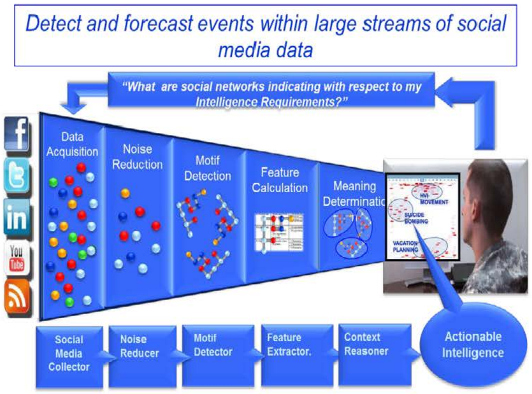
- US Army Research Laboratory (ARL) funded research
- Technology Readiness Level 6 (TRL-6)
- Installed at Aberdeen Proving Ground ARL Laboratory
- Finds and Fingerprints social media users based on interactions:
  - Currently "ISIS", "Business", and Hacker classes (Twitter)


Who is important in the adversary network?
How are they influencing others and vice/versa?
Ego Nets
Ego Net Details

SURF Methodology

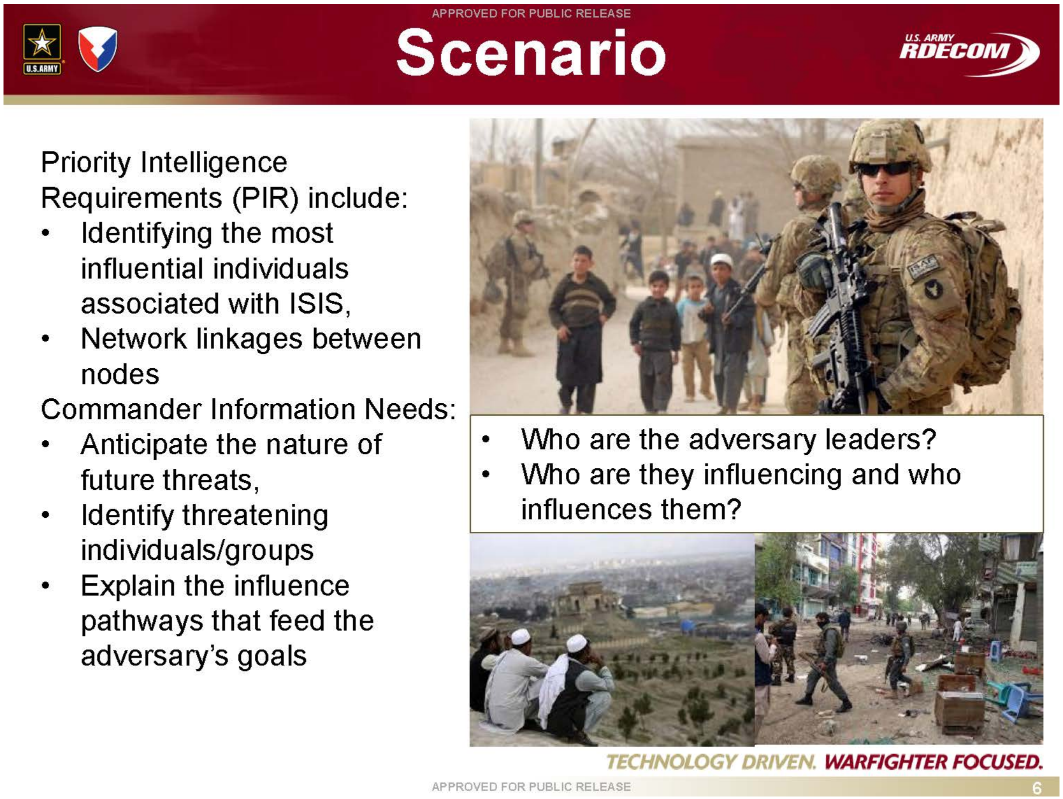
SECURBORATION

Text Motif Discovery



Sociocultural Reasoning Framework (SURF)

- Exploits relationship entities in text to identify potential members of a threat group
- Major components include
  - ➢ Multi-source data acquisition (language agnostic)
  - ➢ Noise reduction
  - ➢ Motif detection based on bio-inspired subgraph identification
  - ➢ Feature extraction to determine threat relationships, levels of threat potential, and group membership



Scenario

Priority Intelligence Requirements (PIR) include:

- Identifying the most influential individuals associated with ISIS,
- Network linkages between nodes

Commander Information Needs:

- Anticipate the nature of future threats,
- Identify threatening individuals/groups
- Explain the influence pathways that feed the adversary's goals

- Who are the adversary leaders?
- Who are they influencing and who influences them?

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Case Study

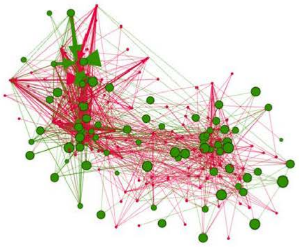
- Three tasks:
  - Analyse the social network,
  - Identify the key influencers, and
  - Analyse the structural elements of message traffic between key influencers
- Data: Twitter feeds
  - 207 nodes (Twitter users)
  - 1,534 edges (friends & followers)

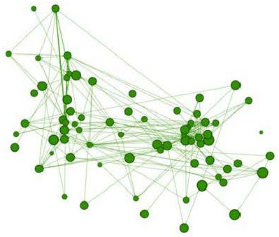
TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Task 1: Analyse the social network

U.S. ARMY RDECOM

- Filter dataset for users identified as ISIS
- Directed graph
  - Green: Ego nodes (predicted ISIS members)
  - Red: Extended Network (friends/followers)



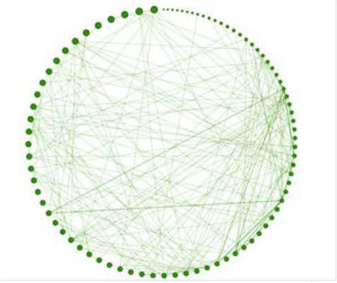


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

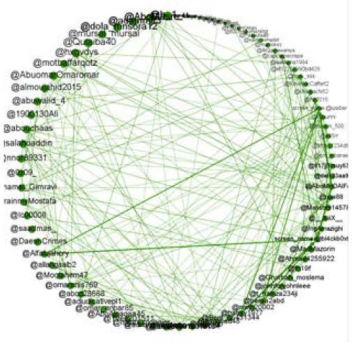
Task 1: Analyse the social network (cont)

U.S. ARMY RDECOM

Filter data with ordered ranking based on eigenvector centrality of each node results in a size-ordered circular layout.



Label graph with Ego Nodes to identify most influential personas. Connections indicate relationships among nodes, with strong connections present at the 2 o'clock position.

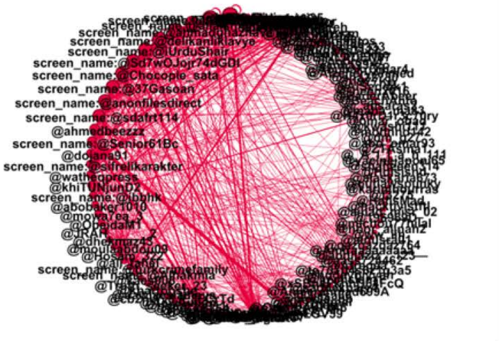


TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Task 2: Identify Who in the Extended Network Is Most Influential

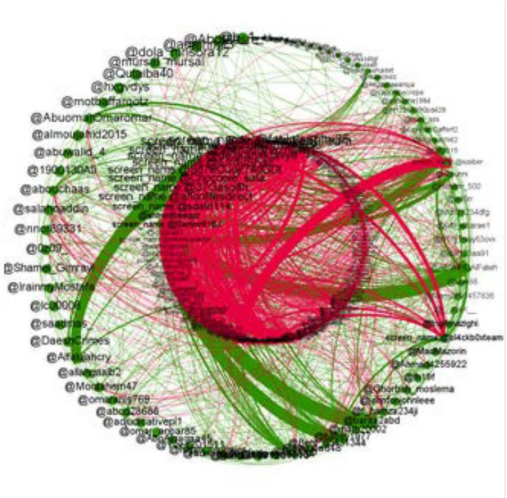
Analysis of friends/followers network

- Highest In-Degree i Extended Network
- These nodes represent the people to whom the ISIS Ego Nodes are listening



Task 2: Identify Who in the Extended Network Is Most Influential

To complete task 2, analyst merges Ego Nodes (green) with Extended Network (red) (F/F) to show relationships between actors and influencers

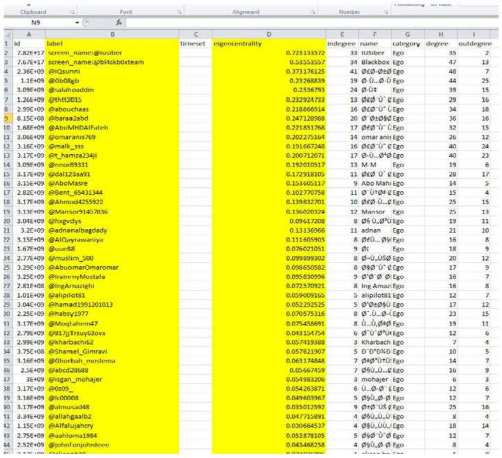


Task 3: Extract Data for Targeting

- The analyst queries the network to show only the Egos and sorts the data (decreasing) on Eigenvector Centrality.
- This yields an ordered list of Ego users who are central to the social network's functionality.
- This list of Twitter handle names is exported to a .csv file.
- Process is repeated for the Extended Network.
- This produces two outcomes:

  1) a list of potential ISIS affiliates and those most important within the network of potential ISIS affiliates, and

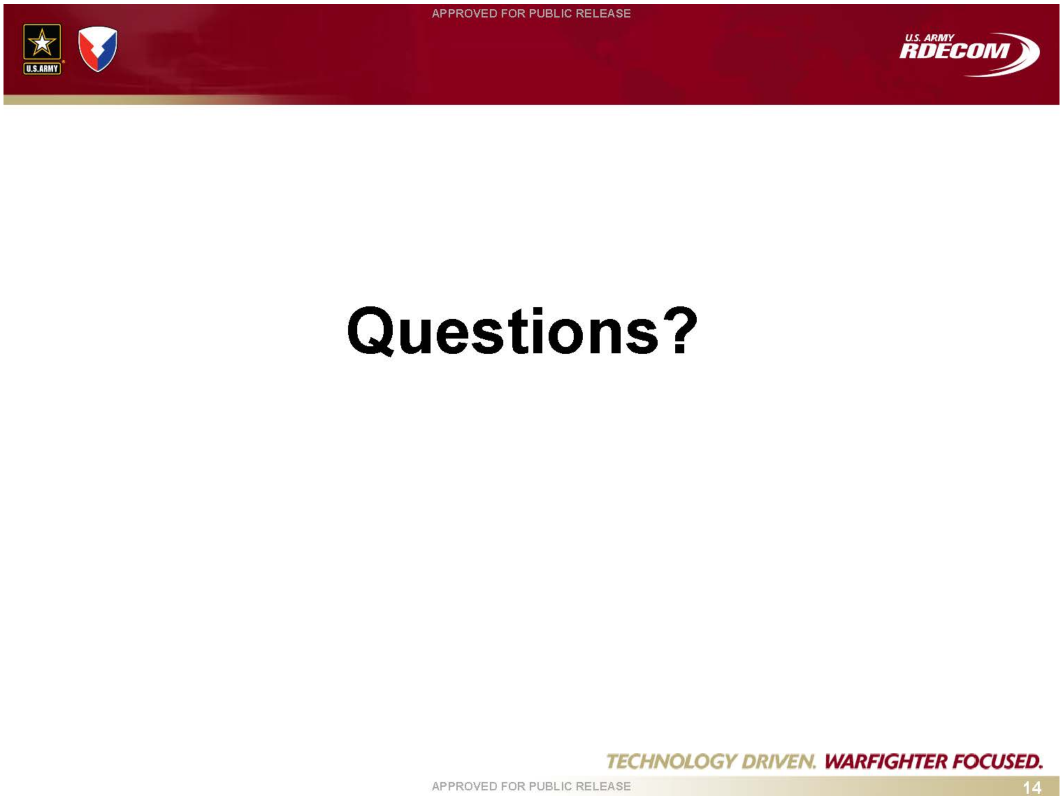
  2) a list of their influencers.



	A	B	C	D	E	F	G	H	I
1	id	label	timeset	eigencentrality	indegree	name	category	degree	outdegree
2	7.82E+17	screen_name:@iusiber		0.723133572	33	IUSiber	Ego	35	2
3	7.67E+17	screen_name:@bl4ckb0xteam		0.53553557	34	Blackbox	Ego	47	13
4	2.36E+09	@IQsunni		0.373176125	41	Ø£Ø-Ø±Ø	Ego	48	7
5	1.1E+09	@Gb08gb		0.23268839	19	Ø-Ù…Ø-Ù	Ego	44	25
6	3.09E+09	@salahoaddin		0.2536793	24	Ø-Ù‡	Ego	39	15
7	1.26E+09	@thtt2015		0.232924733	13	Ø£Ø¨Ù˜ Ø	Ego	29	16
8	2.99E+09	@abouchaas		0.218866914	16	Ø£Ø¨Ù˜ Ù	Ego	34	18
9	8.15E+08	@baraa2abd		0.247128988	20	Ø¨Ø±Ø§Ø	Ego	36	16
10	1.68E+09	@AboMHDAlFateh		0.221851768	17	Ø£Ø¨Ù˜ Ù	Ego	32	15
11	3.06E+09	@omaranis769		0.202275164	14	omar anis	Ego	26	12
12	3.16E+09	@malk_sss		0.191667248	16	Ø£Ø¨Ù˜ Ø	Ego	40	24
13	3.17E+09	@t_hamza234ji		0.200712071	17	Ø-Ù…Ø²Ø	Ego	40	23
14	3.09E+09	@nnor89331		0.192010517	13	M-M	Ego	19	6
15	3.17E+09	@dal123aa91		0.172918105	11	Ø£Ø¨Ù˜ Ø	Ego	28	17
16	3.15E+09	@AboMasre		0.153605117	9	Abo Mah	Ego	14	5
17	2.82E+09	@Bent_65431344		0.102770758	11	Ø¨Ù†Øª Ø	Ego	15	4
18	3.17E+09	@Ahmad4255922		0.139832701	10	Ø£Ø-Ù…Ø	Ego	25	15
19	3.13E+09	@Mansor91457836		0.136020324	12	Mansor	Ego	25	13
20	3.04E+09	@hxgvdys		0.09617208	8	Ø§ Ù„Ø³Ù	Ego	19	11
21	3.2E+09	@adnanalbagdady		0.13136966	11	adnan	Ego	21	10
22	3.15E+09	@AlQayrawaniya		0.111605903	8	Ø£Ù… Ø§	Ego	16	8
23	1.67E+09	@uue88		0.076021051	9	Ø	Ego	18	9
24	2.77E+09	@muslim_500		0.099899302	8	Ø¬Ù„ÙŠØ	Ego	20	12
25	3.29E+09	@AbuomarOmaromar		0.098650582	8	Ø§Ø¨Ù˜ Ø	Ego	17	9
26	3.25E+09	@IrainmyMostafa		0.095830996	9	Ø'Ø¨Ø¨Ø	Ego	16	7
27	2.81E+08	@IngAmazighi		0.072570921	8	Ing Amazi	Ego	16	8
28	1.01E+09	@alipilot81		0.059009165	5	alipilot81	Ego	12	7
29	3.04E+09	@hamad1991201013		0.052292525	5	Ø'Ø±Ø§Ù	Ego	17	12
30	2.25E+09	@habsy1977		0.070575316	8	Ø˜.Ù…Ø-Ù	Ego	23	15
31	3.17E+09	@Moqtahem47		0.075458691	8	Ù…Ù‚ØªØ	Ego	19	11
32	2.79E+09	@817jjTrsuy63ovx		0.043154754	6	Ø˜Ù˜Ø³Ù	Ego	12	6
33	2.99E+09	@kharbachi62		0.057419388	3	Kharbach	Ego	7	4
34	3.75E+08	@Shamel_Gimravi		0.057621907	5	Ð¨Ð°Ð¼Ð	Ego	10	5
35	3.16E+09	@Ghorbah_moslema		0.063174848	7	ØºØ±Ù†Ù	Ego	14	7
36	2.5E+09	@abcd28688		0.05667459	7	Ø§Ù„Ù…Ø	Ego	16	9
37	3E+09	@isgan_mohajer		0.054983206	3	mohajer	Ego	6	3
38	3.17E+09	@0z09_		0.054263871	6	Ù…Ø-Ø˜ Ø	Ego	12	6
39	3.16E+09	@lc00008		0.049603967	5	Ø§Ù„Ø-Ø	Ego	12	7
40	3.17E+09	@almusad48		0.035012592	9	Ø±Ø¨ÙŠ Ø	Ego	25	16
41	3.34E+09	@allahgaalb2		0.047715891	4	Ø§Ù„Ù„Ù	Ego	8	4
42	1.15E+09	@Alfalujahcry		0.030664537	4	Ø§Ù„ÙÙ„Ù	Ego	18	14
43	2.75E+09	@aahloma1984		0.052878105	5	Ø§Ø¨Ù˜Ø	Ego	12	7
44	2.52E+09	@johnfonjohnleee		0.043468258	4	Ø§Ù„Ø-Ø	Ego	8	4

Conclusions

- Automated discovery and monitoring of targeted user classes (e.g. ISIS sympathizer) from social media text, regardless of language, provided an accurate and timely way to identify threat groups that will reduce cognitive workload and mental fatigue for analysts.
- Pilot tests with operational analysts indicate SURF saves analysts an estimated 80-85% in analytic processing time.
- Analysts can create tailored watchlists based on the social networks of those classified as likely ISIS members.



APPROVED FOR PUBLIC RELEASE

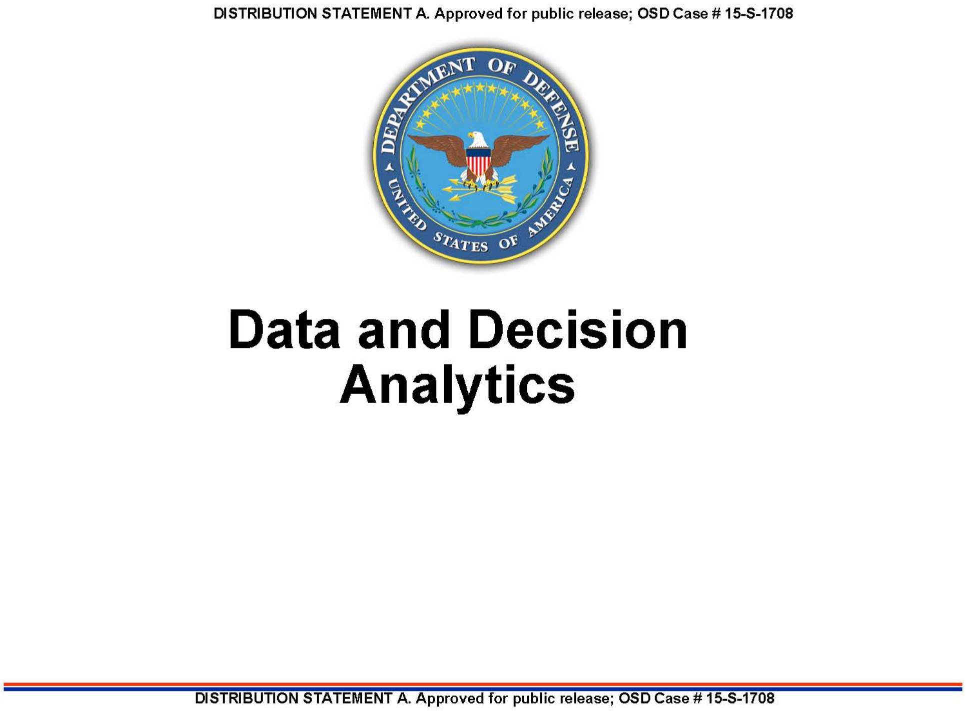
U.S. ARMY RDECOM

Questions?

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

APPROVED FOR PUBLIC RELEASE

14



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Data and Decision Analytics

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Data Analytics Assessment Overview

Problem: There is currently no standard way to implement and assess performance for data analytics

- Heterogeneous data sources/algorithms without ground truth
- Hard to know what capability is being purchased with few means to assess performance of service
- Dynamic mission space with changing requirements

Solution: Data analytics framework

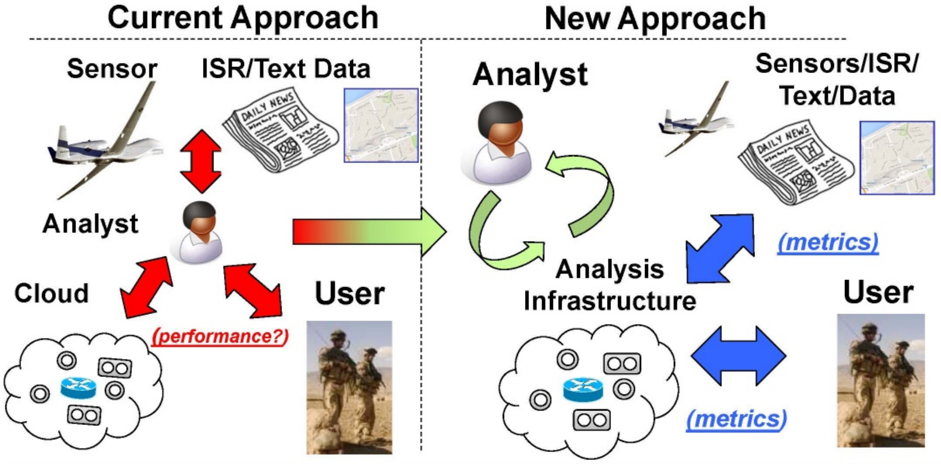
- Standard data models with ground truth
- Development framework to standardize risk analytics on information sources, algorithms, and processing
- Adaptable framework that can change as mission requirements change

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

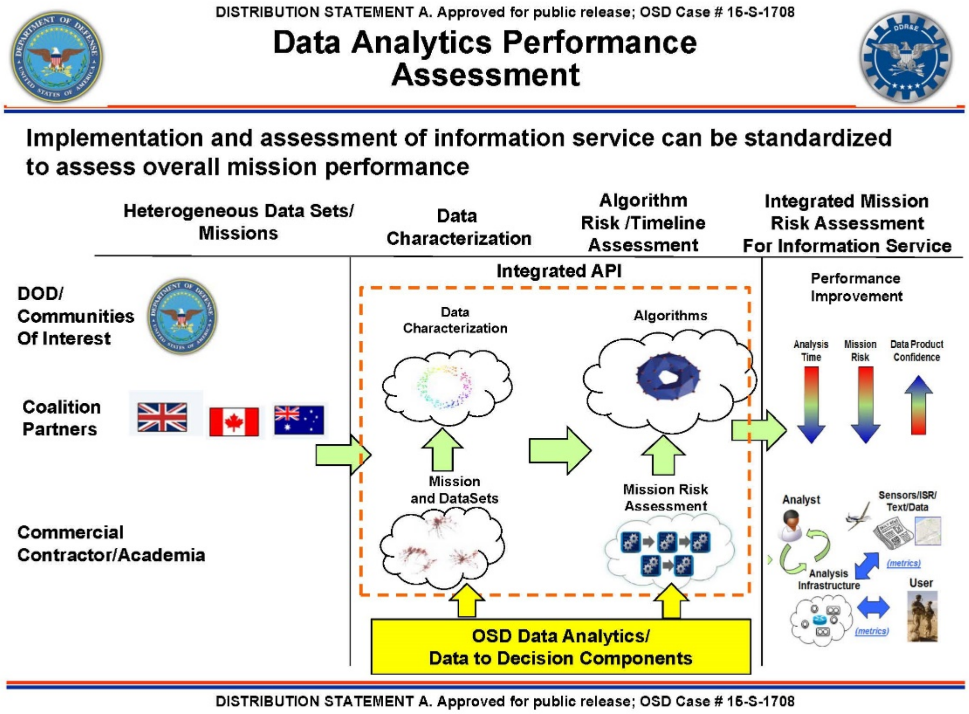
DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

D2D/Data Analytics Approach

Analyst oversees delivery of information products to customer with rigorous quality of service guarantees



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Data Analytics Performance Assessment

Implementation and assessment of information service can be standardized to assess overall mission performance

Heterogeneous Data Sets/ Missions

Data Characterization

Algorithm Risk /Timeline Assessment

Integrated Mission Risk Assessment For Information Service

DOD/ Communities Of Interest

Coalition Partners

Commercial Contractor/Academia

Integrated API

Data Characterization

Algorithms

Mission and DataSets

Mission Risk Assessment

OSD Data Analytics/ Data to Decision Components

Performance Improvement

Analysis Time

Mission Risk

Data Product Confidence

Analyst

Sensors/ISR/ Text/Data

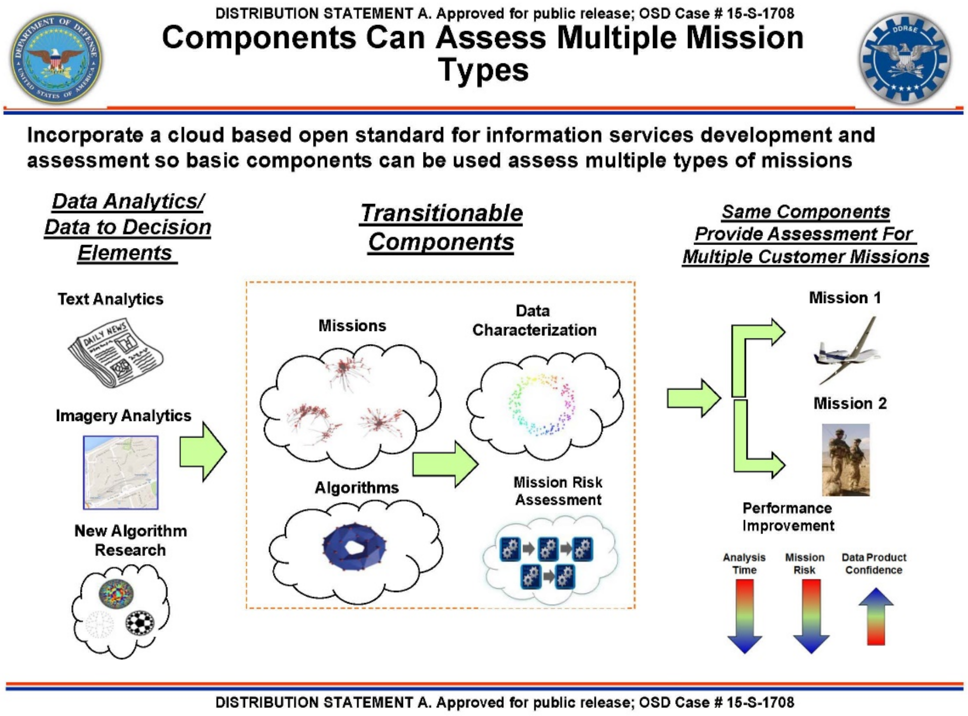
(metrics)

Analysis Infrastructure

User

(metrics)

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Components Can Assess Multiple Mission Types

Incorporate a cloud based open standard for information services development and assessment so basic components can be used assess multiple types of missions

Data Analytics/ Data to Decision Elements

Transitionable Components

Same Components Provide Assessment For Multiple Customer Missions

Text Analytics

Imagery Analytics

New Algorithm Research

Missions

Data Characterization

Algorithms

Mission Risk Assessment

Mission 1

Mission 2

Performance Improvement

Analysis Time

Mission Risk

Data Product Confidence

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Transition Models

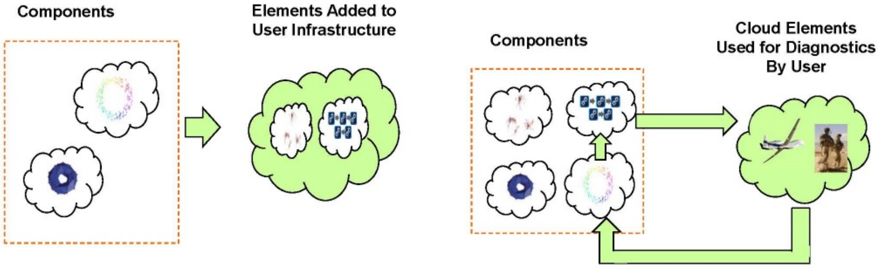
Models can either be added to existing infrastructure or used by existing infrastructure as diagnostics for performance

Model 1

(direct integration of components)

Model 2

(user integrates remote elements for their analysis)



Components

Elements Added to User Infrastructure

Components

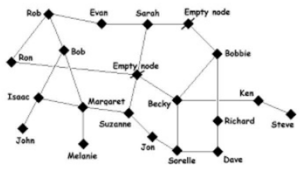
Cloud Elements Used for Diagnostics By User

Mission and Data Set Components

Standard threat or mission graphs and the associated data needed to assess a particular threat are available for baseline assessment and design of future missions analysis

Standard Mission Graphs

Scenario Graph Specifies What Data Should Be Collected

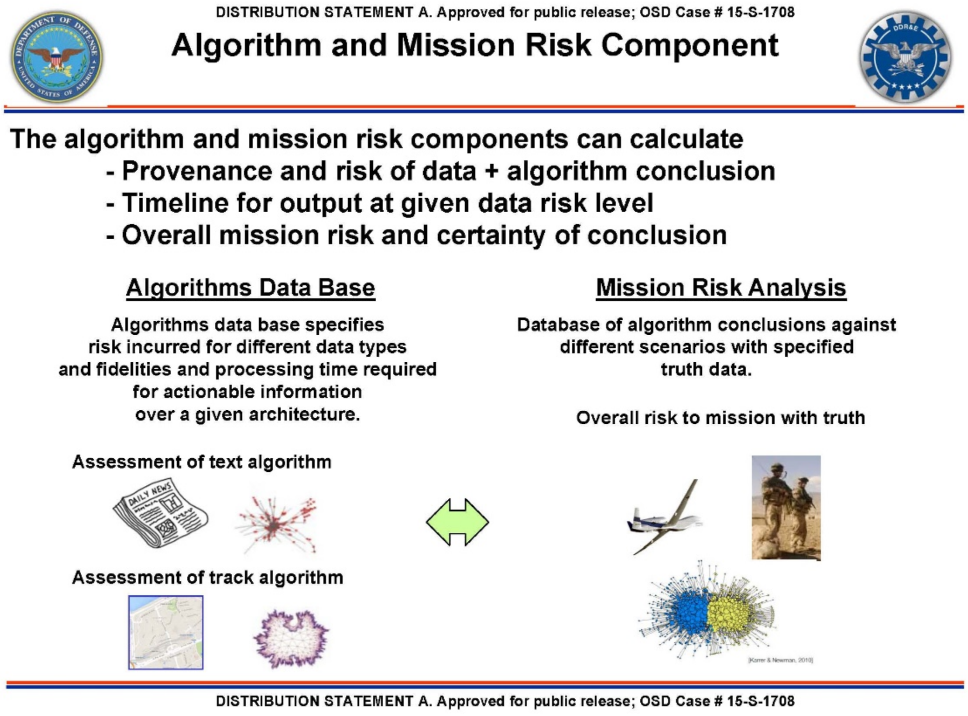


Standard Data Sets

Standard Data Sets Specify Ground Truth for Different Data Types & Provenance of Relevant Data

Imagery Truth Data

Text Analytic Data



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Algorithm and Mission Risk Component

The algorithm and mission risk components can calculate

- Provenance and risk of data + algorithm conclusion
- Timeline for output at given data risk level
- Overall mission risk and certainty of conclusion

Algorithms Data Base

Algorithms data base specifies risk incurred for different data types and fidelities and processing time required for actionable information over a given architecture.

Assessment of text algorithm

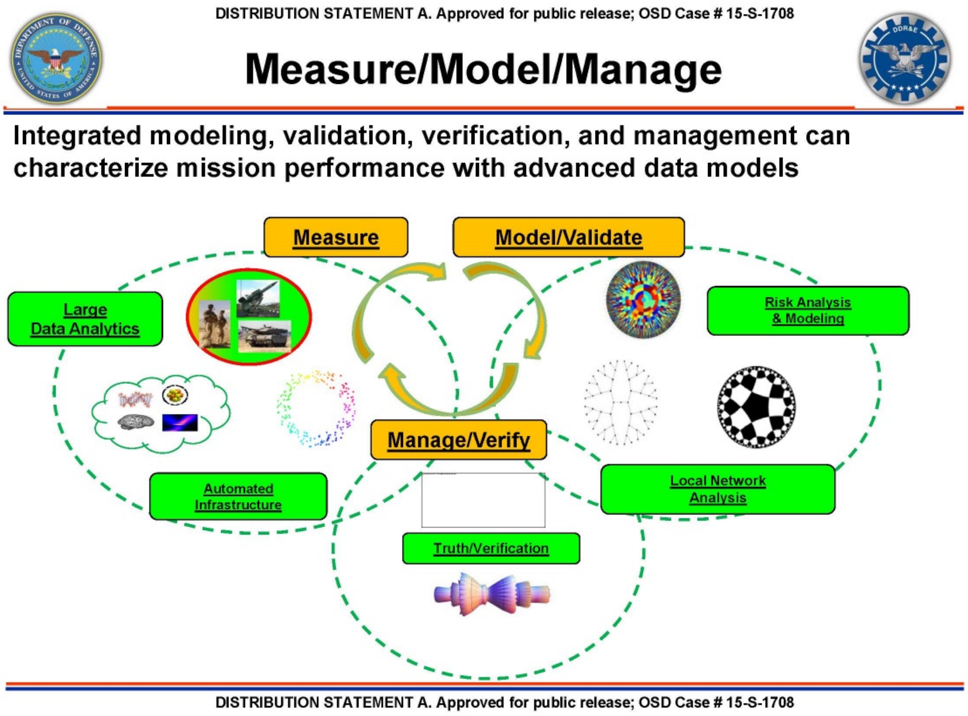
Assessment of track algorithm

Mission Risk Analysis

Database of algorithm conclusions against different scenarios with specified truth data.

Overall risk to mission with truth

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Measure/Model/Manage

Integrated modeling, validation, verification, and management can characterize mission performance with advanced data models

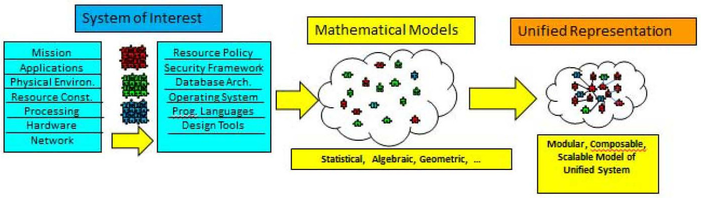
DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Measurement

We wish to understand how to measure the state of a mission on an infrastructure

What to measure?



System of Interest

Mission
Applications
Physical Environ.
Resource Const.
Processing
Hardware
Network

Resource Policy
Security Framework
Database Arch.
Operating System
Prog. Languages
Design Tools

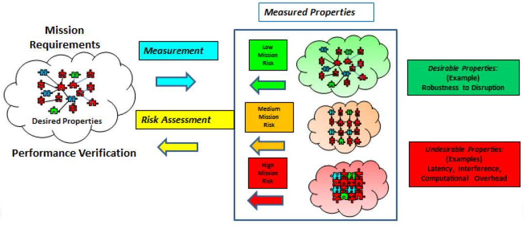
Mathematical Models

Statistical, Algebraic, Geometric, ...

Unified Representation

Modular, Composable, Scalable Model of Unified System

How to measure?



Mission Requirements

Desired Properties

Performance Verification

Measurement

Risk Assessment

Measured Properties

Low Mission Risk

Medium Mission Risk

High Mission Risk

Desirable Properties:
(Example)
Robustness to Disruption

Undesirable Properties:
(Examples)
Latency, Interference, Computational Overhead

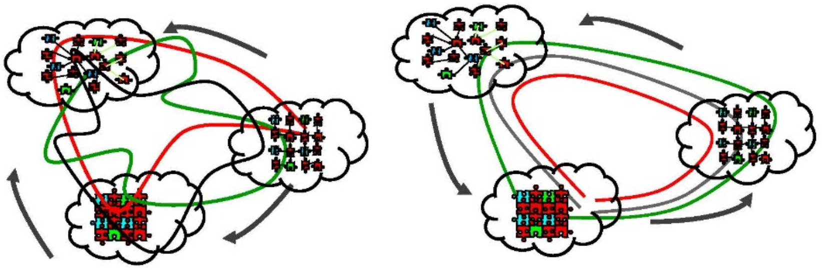
DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Modeling

We must have validated models of mission performance which can come from known models or empirical data

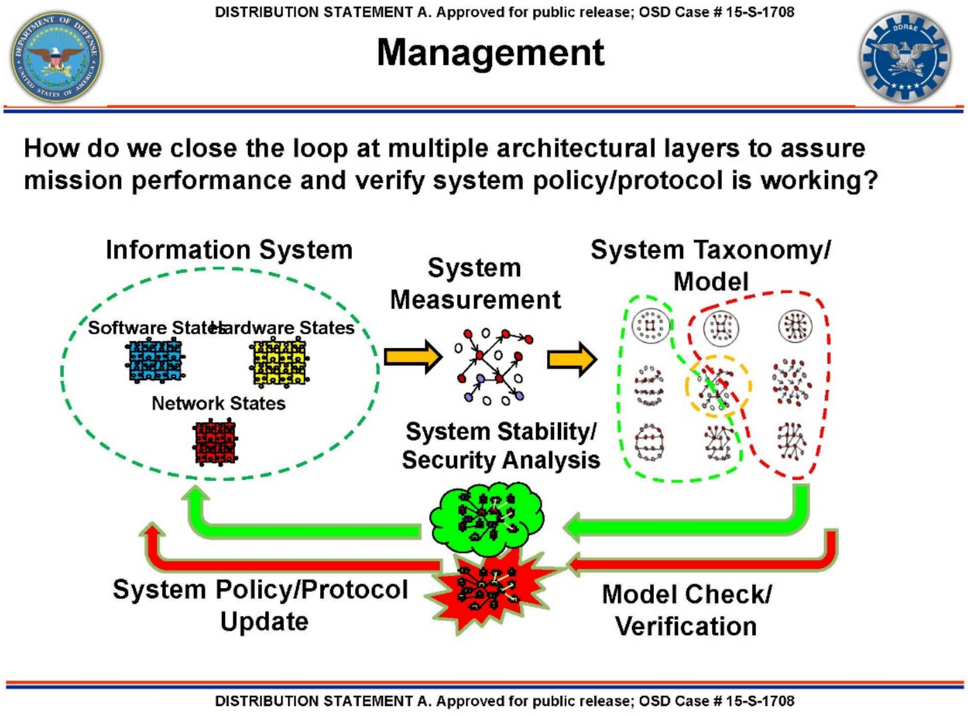
Mission Operation Trade-space



Un-validated Modalities (high mission risk)

Validated Modalities (low mission risk)

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

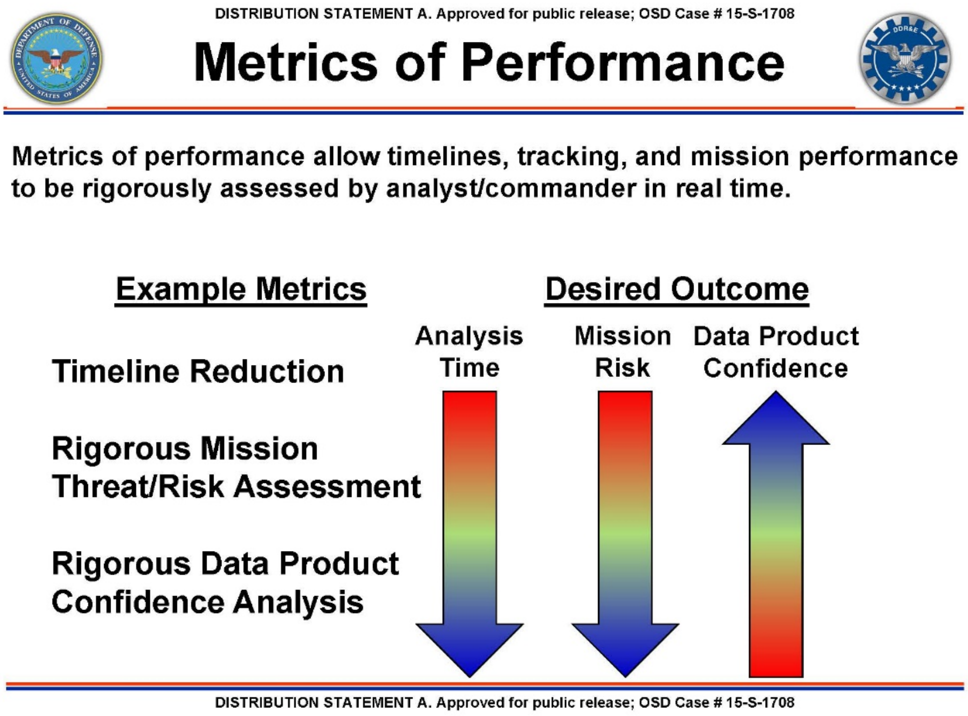


DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Management

How do we close the loop at multiple architectural layers to assure mission performance and verify system policy/protocol is working?

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Metrics of Performance

Metrics of performance allow timelines, tracking, and mission performance to be rigorously assessed by analyst/commander in real time.

Example Metrics

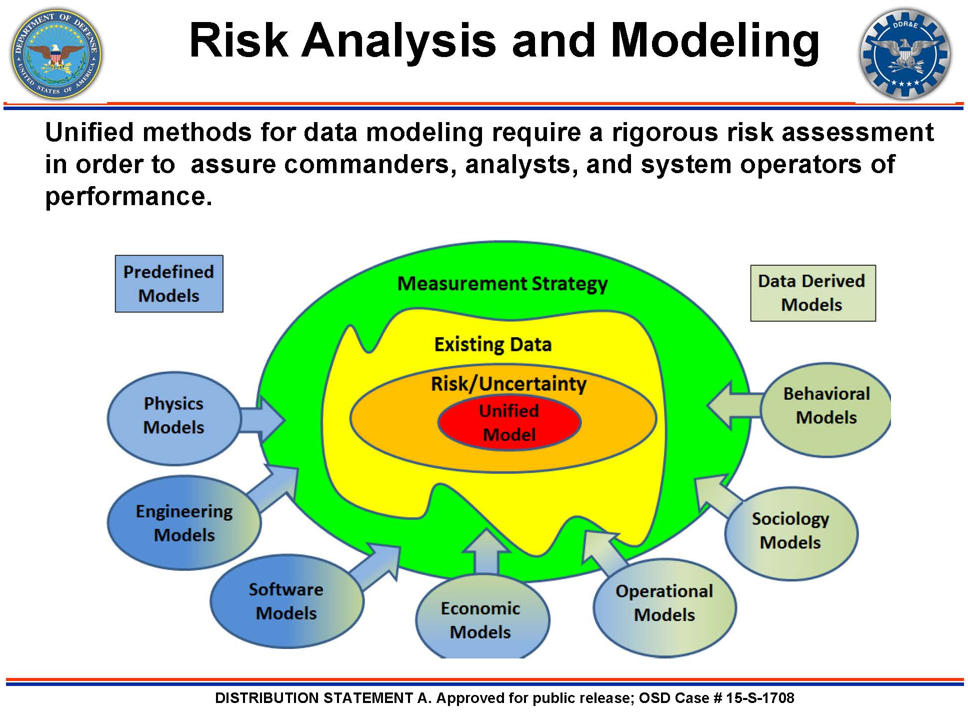
Timeline Reduction

Rigorous Mission Threat/Risk Assessment

Rigorous Data Product Confidence Analysis

Desired Outcome

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



Risk Analysis and Modeling

Unified methods for data modeling require a rigorous risk assessment in order to assure commanders, analysts, and system operators of performance.

Predefined Models

Data Derived Models

Measurement Strategy

Existing Data

Risk/Uncertainty

Unified Model

Physics Models

Engineering Models

Software Models

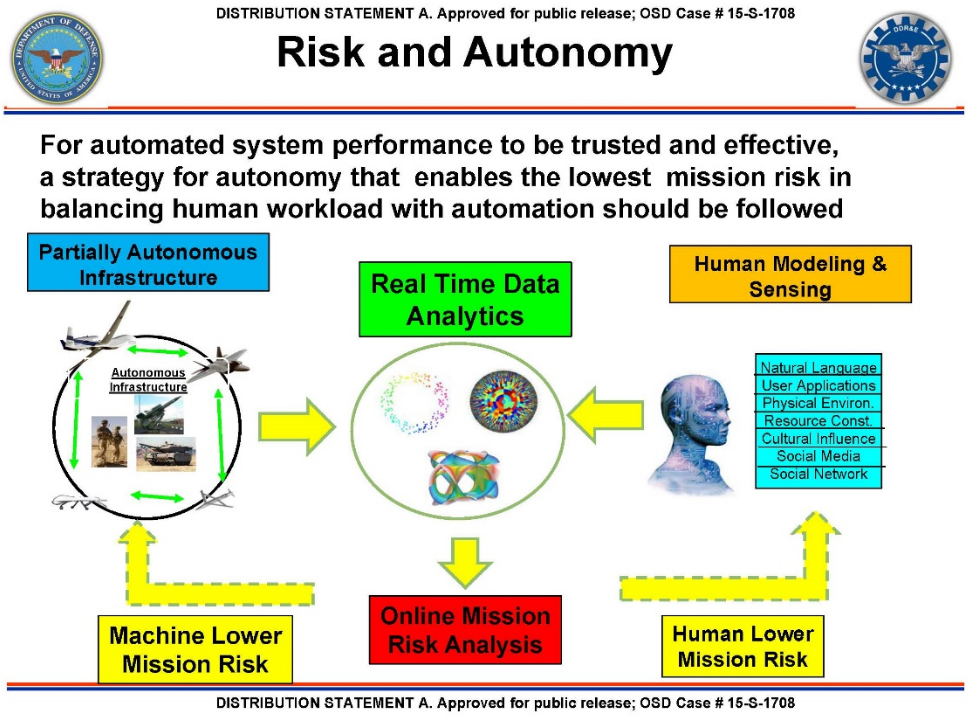
Economic Models

Operational Models

Sociology Models

Behavioral Models

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Risk and Autonomy

For automated system performance to be trusted and effective, a strategy for autonomy that enables the lowest mission risk in balancing human workload with automation should be followed

Partially Autonomous Infrastructure

Real Time Data Analytics

Human Modeling & Sensing

Autonomous Infrastructure

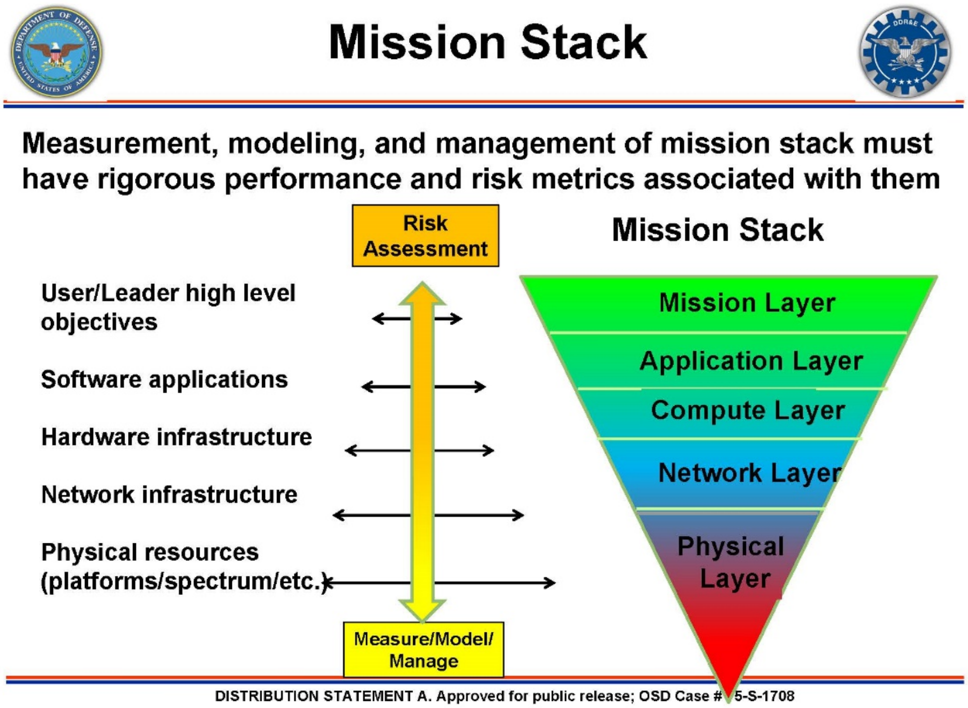
Natural Language
User Applications
Physical Environ.
Resource Const.
Cultural Influence
Social Media
Social Network

Machine Lower Mission Risk

Online Mission Risk Analysis

Human Lower Mission Risk

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



Mission Stack

Measurement, modeling, and management of mission stack must have rigorous performance and risk metrics associated with them

Risk Assessment

Mission Stack

User/Leader high level objectives

Software applications

Hardware infrastructure

Network infrastructure

Physical resources (platforms/spectrum/etc.)

Measure/Model/ Manage

Mission Layer

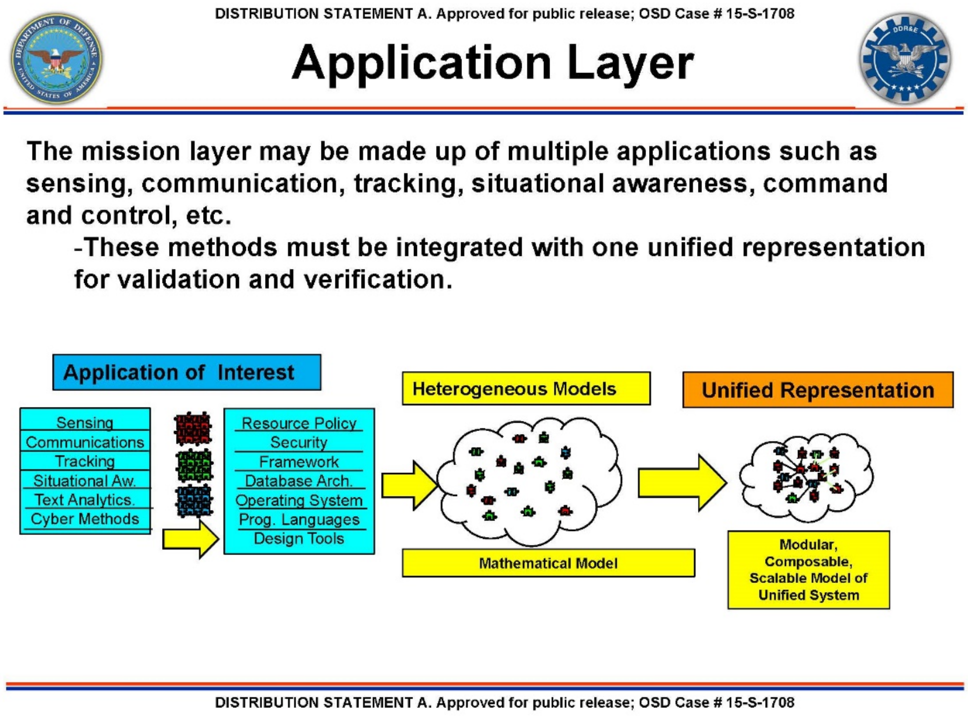
Application Layer

Compute Layer

Network Layer

Physical Layer

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Application Layer

The mission layer may be made up of multiple applications such as sensing, communication, tracking, situational awareness, command and control, etc.

-These methods must be integrated with one unified representation for validation and verification.

Application of Interest

Sensing
Communications
Tracking
Situational Aw.
Text Analytics
Cyber Methods

Resource Policy
Security
Framework
Database Arch.
Operating System
Prog. Languages
Design Tools

Heterogeneous Models

Mathematical Model

Unified Representation

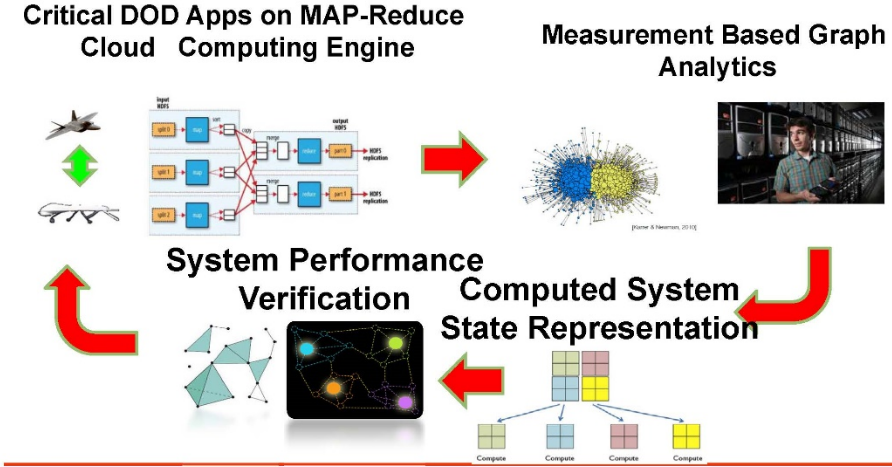
Modular, Composable, Scalable Model of Unified System

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Compute Layer

Current computational infrastructures (cloud resources) are currently highly distributed and resource allocation is static. Making this process more dynamic will create resilient system performance.



Critical DOD Apps on MAP-Reduce Cloud Computing Engine

Measurement Based Graph Analytics

System Performance Verification

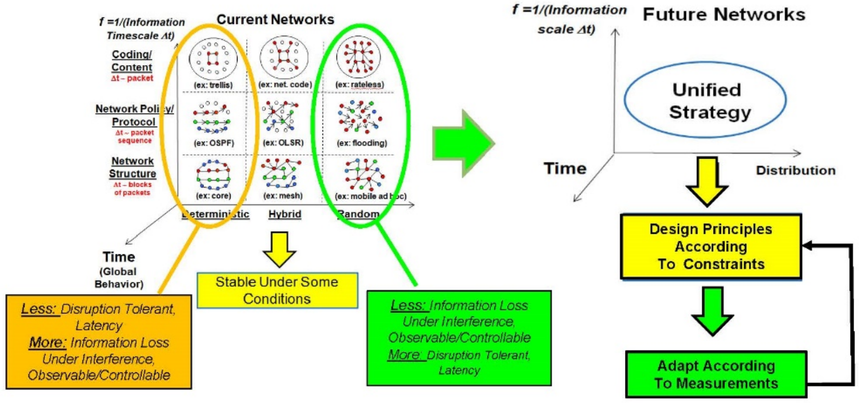
Computed System State Representation

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Network Layer

Advances such as software defined networks are changing stove piped network management to a heterogeneous management problem which requires dynamic assessment



Current Networks

Less: Disruption Tolerant, Latency
More: Information Loss Under Interference, Observable/Controllable

Stable Under Some Conditions

Less: Information Loss Under Interference, Observable/Controllable
More: Disruption Tolerant, Latency

Future Networks

Unified Strategy

Design Principles According To Constraints

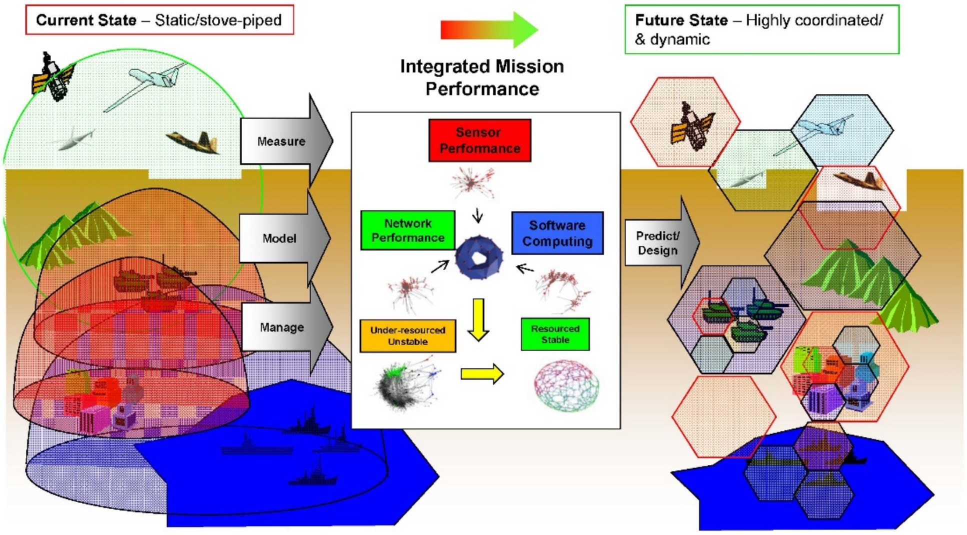
Adapt According To Measurements

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Physical Layer

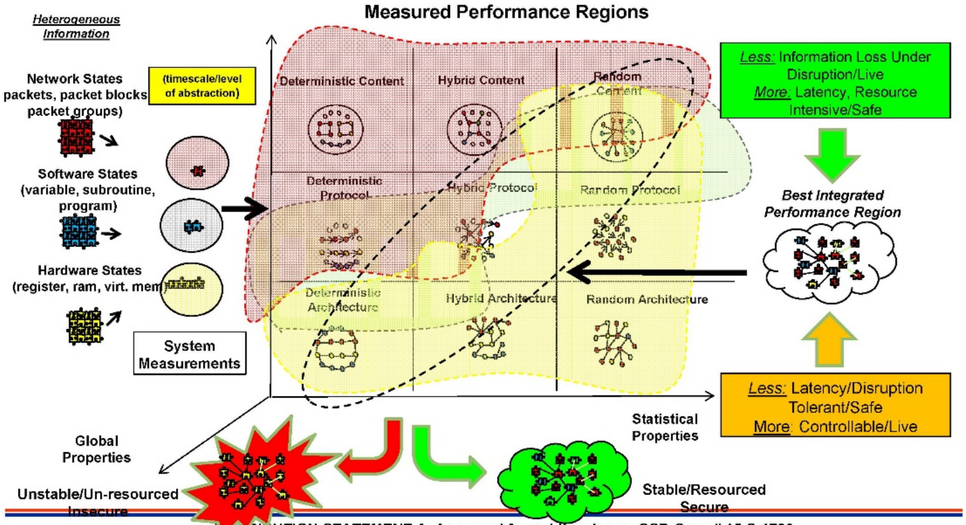
Commercial pressure on spectrum is changing the static and highly segregated assumptions about physical layer performance.



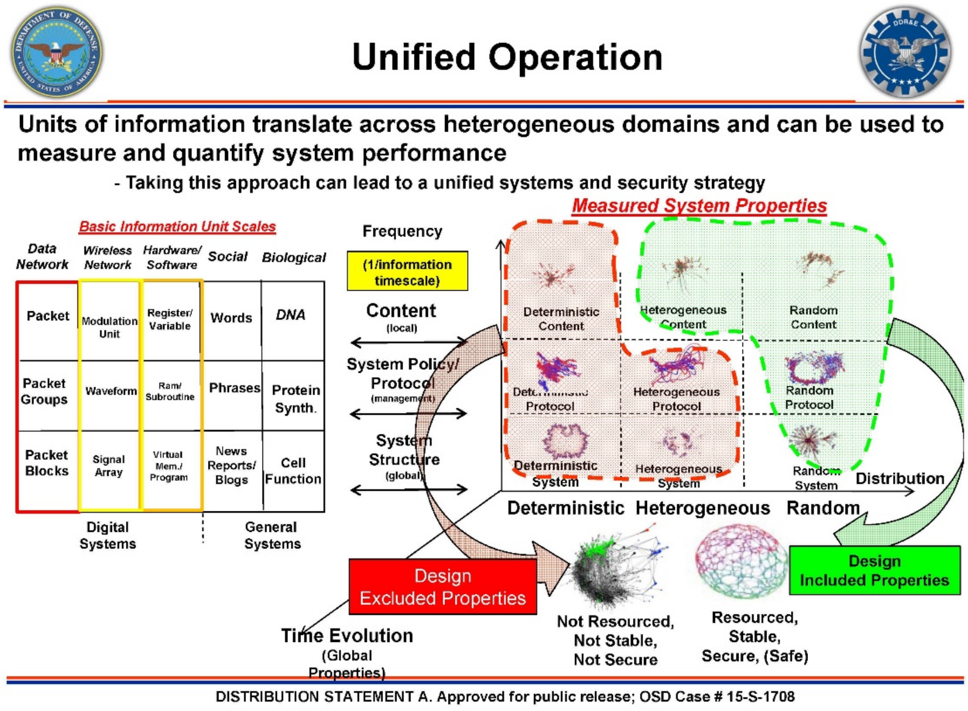
DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Unified Operation

Measure and verify information system properties among various system constraints


Measured Performance Regions
Heterogeneous Information
Network States packets, packet blocks packet groups)
(timescale/level of abstraction)
Software States (variable, subroutine, program)
Hardware States (register, ram, virt. mem)
System Measurements
Deterministic Content
Hybrid Content
Random Content
Deterministic Protocol
Hybrid Protocol
Random Protocol
Deterministic Architecture
Hybrid Architecture
Random Architecture
Less: Information Loss Under Disruption/Live
More: Latency, Resource Intensive/Safe
Best Integrated Performance Region
Less: Latency/Disruption Tolerant/Safe
More: Controllable/Live
Statistical Properties
Global Properties
Unstable/Un-resourced Insecure
Stable/Resourced Secure

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



Unified Operation

Units of information translate across heterogeneous domains and can be used to measure and quantify system performance

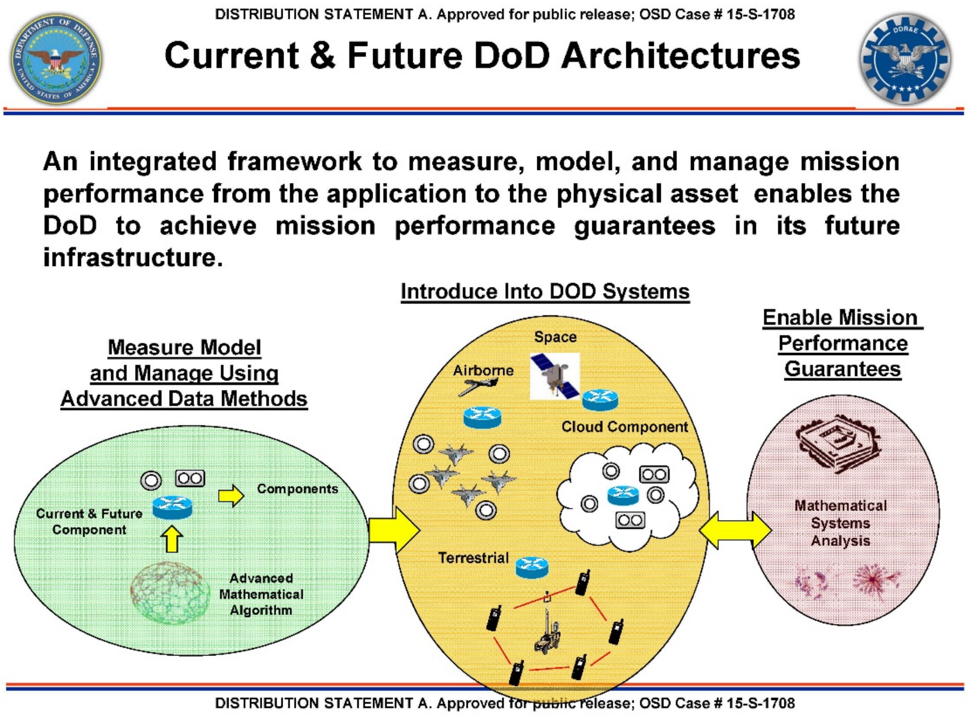
- Taking this approach can lead to a unified systems and security strategy

Basic Information Unit Scales

Data Network	*Wireless Network*	*Hardware/ Software*	*Social*	*Biological*
Packet	Modulation Unit	Register/ Variable	Words	*DNA*
Packet Groups	Waveform	Ram/ Subroutine	Phrases	Protein Synth.
Packet Blocks	Signal Array	Virtual Mem./ Program	News Reports/ Blogs	Cell Function

Digital Systems | General Systems

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



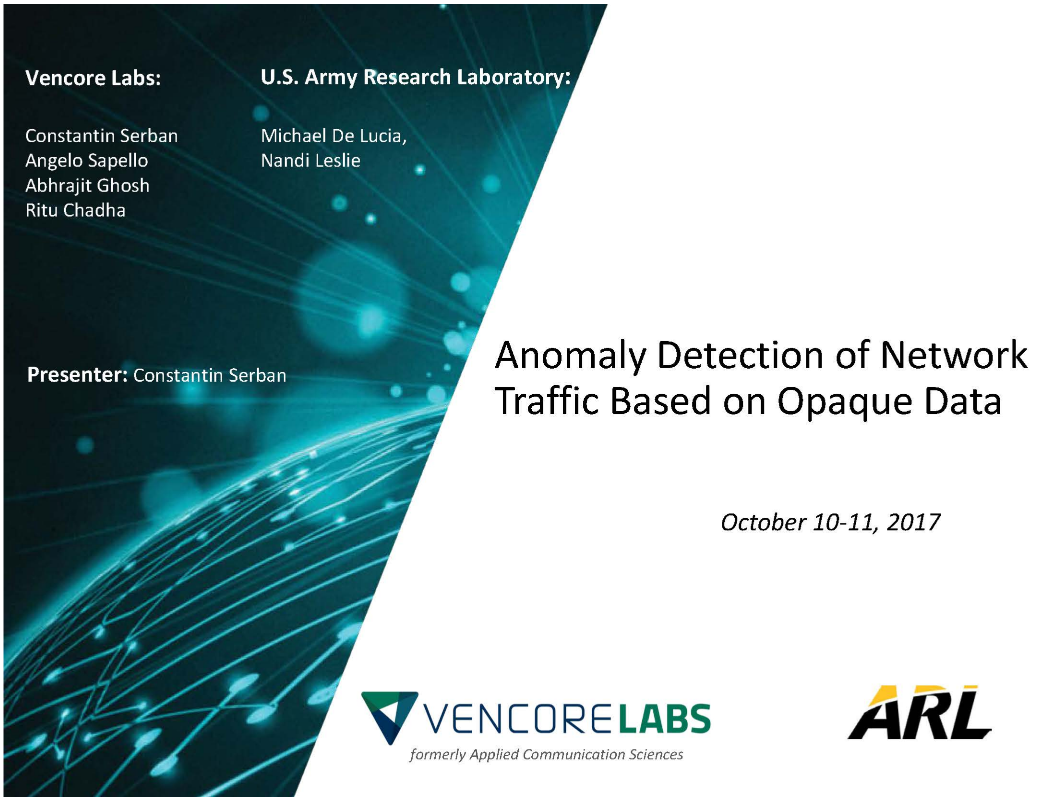
DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708

Current & Future DoD Architectures

An integrated framework to measure, model, and manage mission performance from the application to the physical asset enables the DoD to achieve mission performance guarantees in its future infrastructure.

Introduce Into DOD Systems
Measure Model and Manage Using Advanced Data Methods
Enable Mission Performance Guarantees
Current & Future Component
Components
Advanced Mathematical Algorithm
Space
Airborne
Cloud Component
Terrestrial
Mathematical Systems Analysis

DISTRIBUTION STATEMENT A. Approved for public release; OSD Case # 15-S-1708



Vencore Labs:

Constantin Serban
Angelo Sapello
Abhrajit Ghosh
Ritu Chadha

U.S. Army Research Laboratory:

Michael De Lucia,
Nandi Leslie

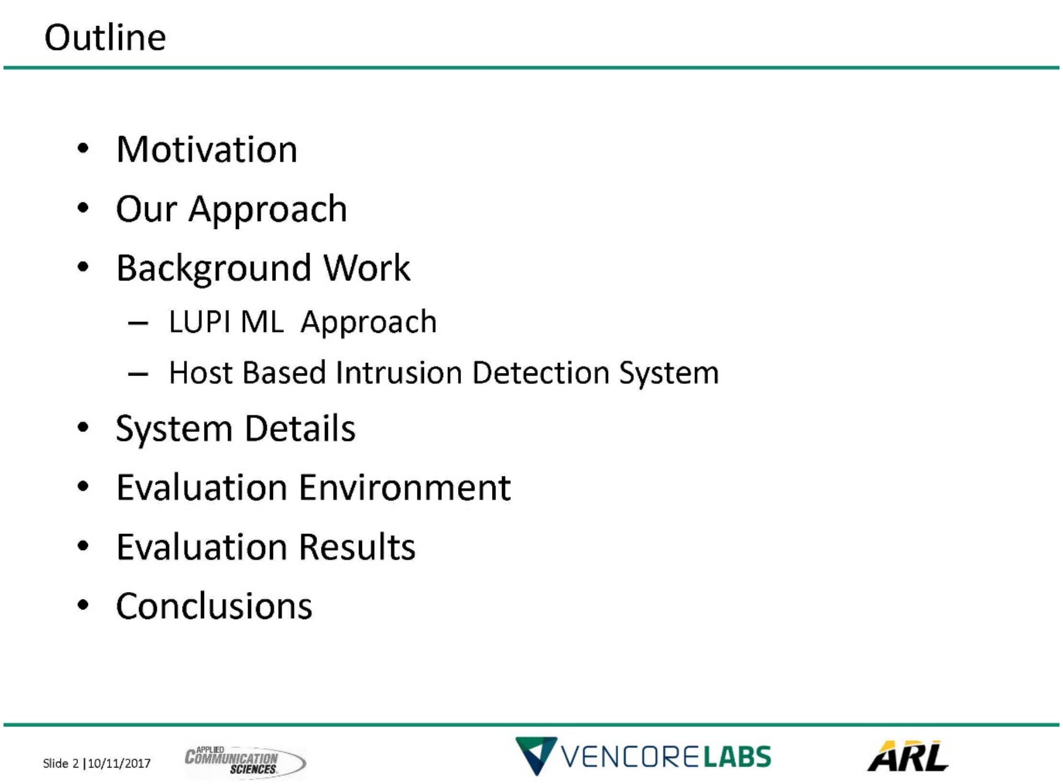
Presenter: Constantin Serban

Anomaly Detection of Network Traffic Based on Opaque Data

October 10-11, 2017

VENCORELABS
formerly Applied Communication Sciences

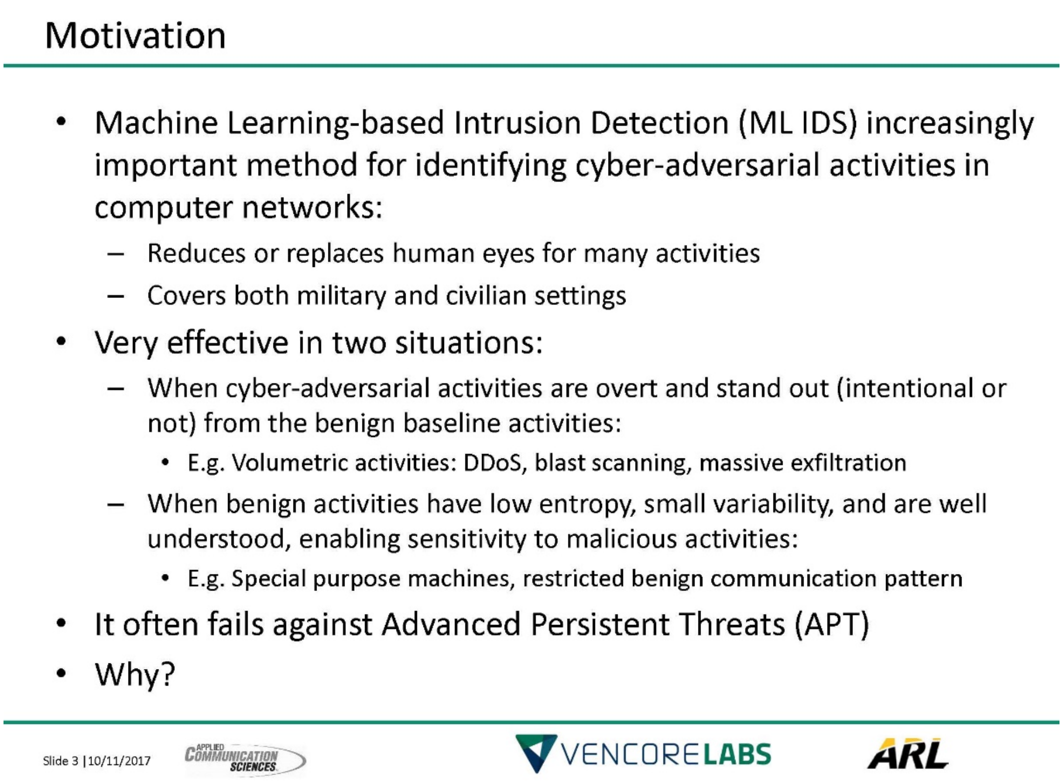
ARL



Outline

- Motivation
- Our Approach
- Background Work
  – LUPI ML Approach
  – Host Based Intrusion Detection System
- System Details
- Evaluation Environment
- Evaluation Results
- Conclusions

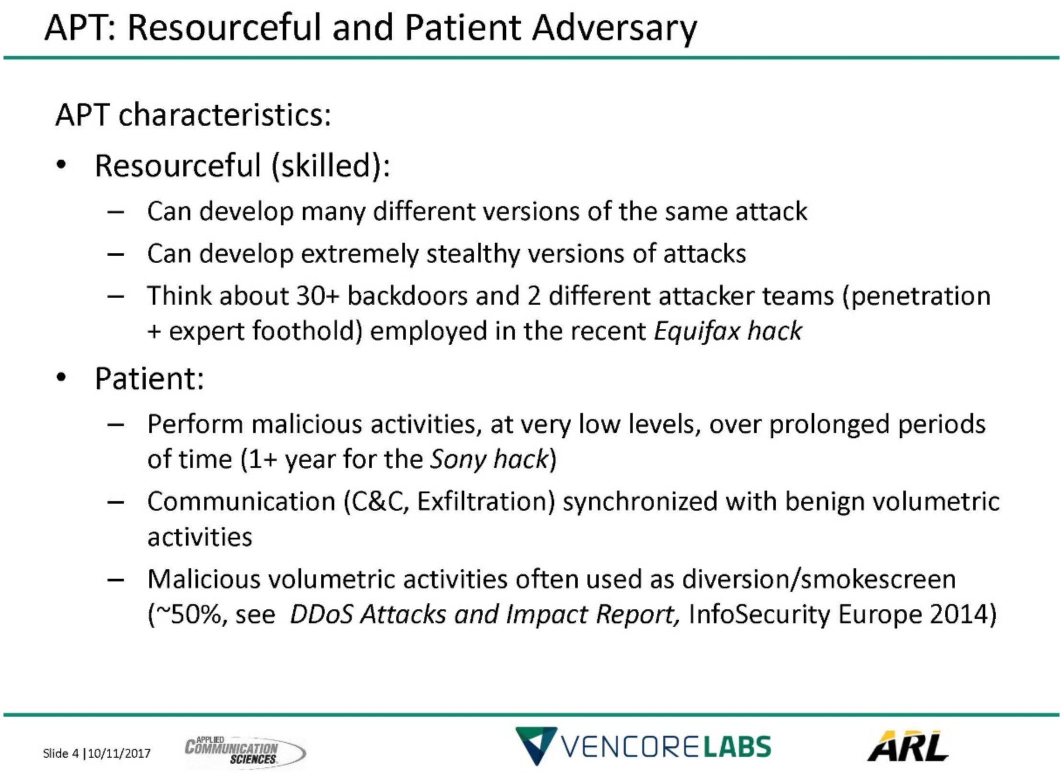
Slide 2 |10/11/2017



Motivation

- Machine Learning-based Intrusion Detection (ML IDS) increasingly important method for identifying cyber-adversarial activities in computer networks:
  - Reduces or replaces human eyes for many activities
  - Covers both military and civilian settings
- Very effective in two situations:
  - When cyber-adversarial activities are overt and stand out (intentional or not) from the benign baseline activities:
    - E.g. Volumetric activities: DDoS, blast scanning, massive exfiltration
  - When benign activities have low entropy, small variability, and are well understood, enabling sensitivity to malicious activities:
    - E.g. Special purpose machines, restricted benign communication pattern
- It often fails against Advanced Persistent Threats (APT)
- Why?

Slide 3 |10/11/2017



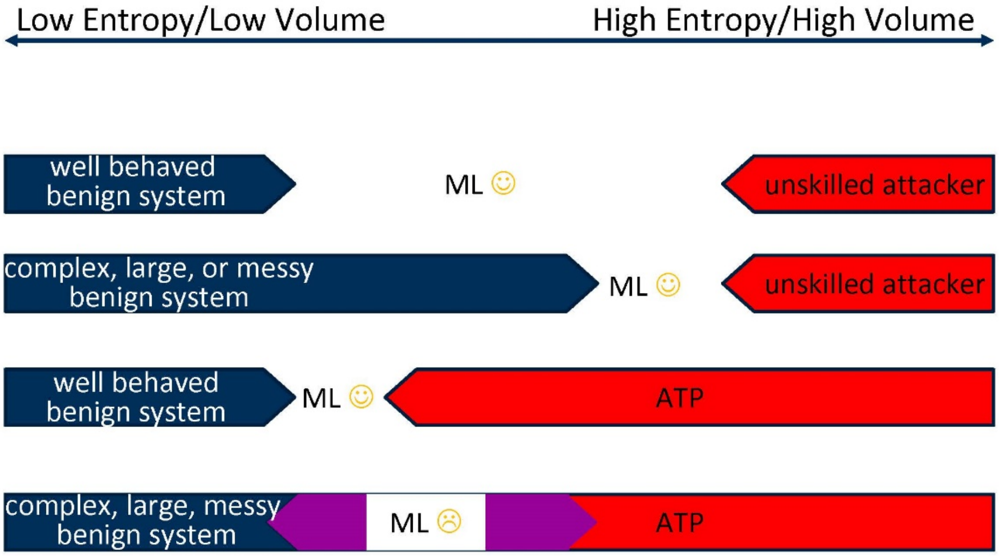
APT: Resourceful and Patient Adversary

APT characteristics:

- Resourceful (skilled):
  - Can develop many different versions of the same attack
  - Can develop extremely stealthy versions of attacks
  - Think about 30+ backdoors and 2 different attacker teams (penetration + expert foothold) employed in the recent *Equifax hack*
- Patient:
  - Perform malicious activities, at very low levels, over prolonged periods of time (1+ year for the *Sony hack*)
  - Communication (C&C, Exfiltration) synchronized with benign volumetric activities
  - Malicious volumetric activities often used as diversion/smokescreen (~50%, see *DDoS Attacks and Impact Report,* InfoSecurity Europe 2014)

Slide 4 |10/11/2017

APT vs. “*Your Daily Hack*”



Low Entropy/Low Volume

High Entropy/High Volume

well behaved benign system

ML ☺

unskilled attacker

complex, large, or messy benign system

ML ☺

unskilled attacker

well behaved benign system

ML ☺

ATP

complex, large, messy benign system

ML ☹

ATP

ML: Machine Learning Model

APT vs. “*Your Daily Hack*” (cont’ed)

Hopeless?

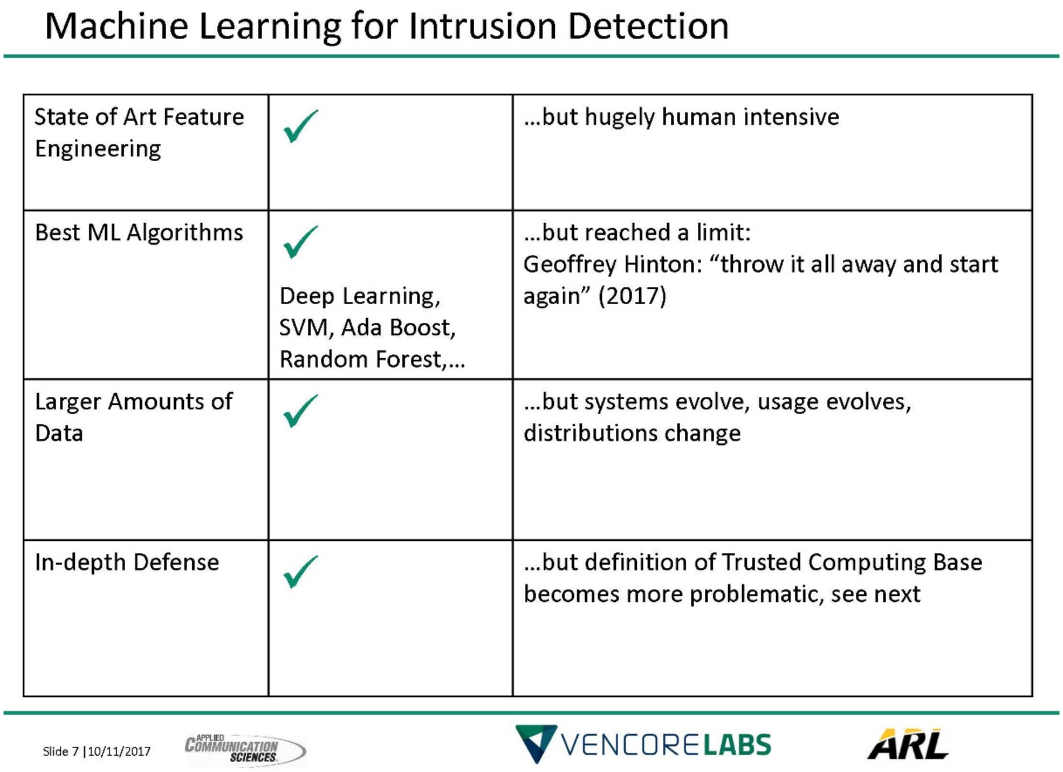
- Do not have control over attacker skill or patience:
  - Must assume APT
- Do not have (full) control over the defended system:
  - Function takes precedence
  - Mission takes precedence

What to do:

- We must understand (i.e. model) much more accurately the System Under Defense (SUD)
  - Reduce the entropy of the SUD (apparent to ML)
- …build better models!

How to do it:

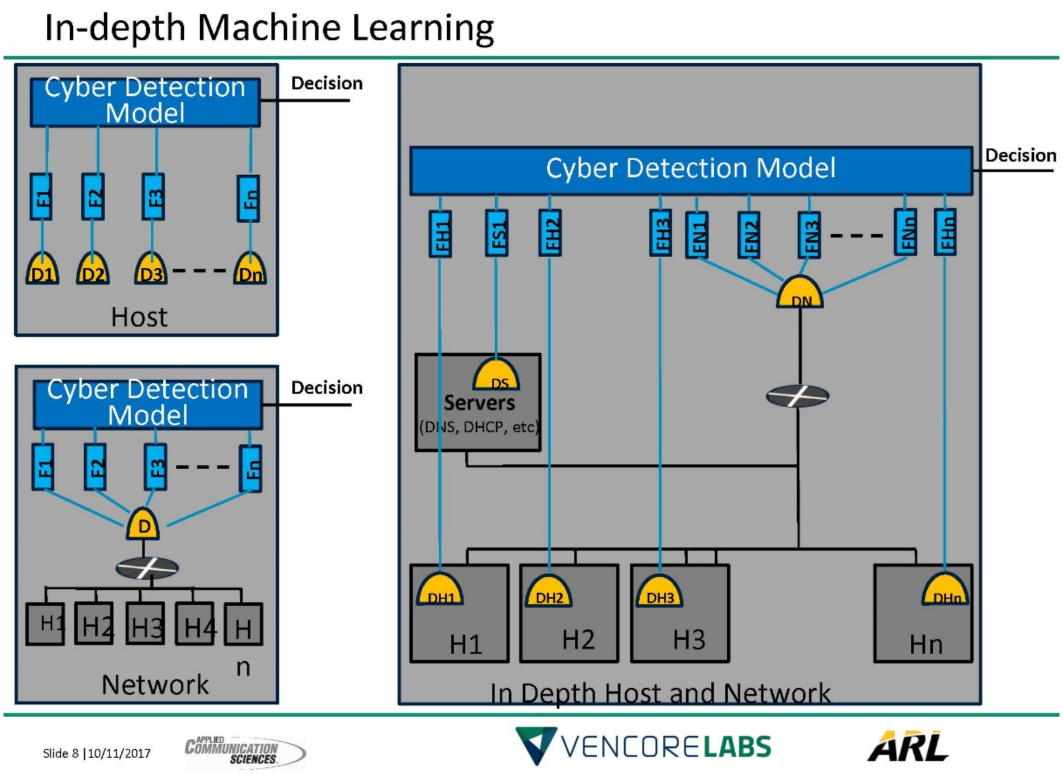
- State-of-the-art *feature* engineering
- Best ML algorithms
- Large amounts of data for developing models
- *In-depth defense*



Machine Learning for Intrusion Detection

State of Art Feature Engineering	✓	…but hugely human intensive
Best ML Algorithms	✓ Deep Learning, SVM, Ada Boost, Random Forest,…	…but reached a limit: Geoffrey Hinton: “throw it all away and start again” (2017)
Larger Amounts of Data	✓	…but systems evolve, usage evolves, distributions change
In-depth Defense	✓	…but definition of Trusted Computing Base becomes more problematic, see next

Slide 7 |10/11/2017



In-depth Machine Learning

Slide 8 |10/11/2017

Weaknesses of In-depth Machine Learning IDS

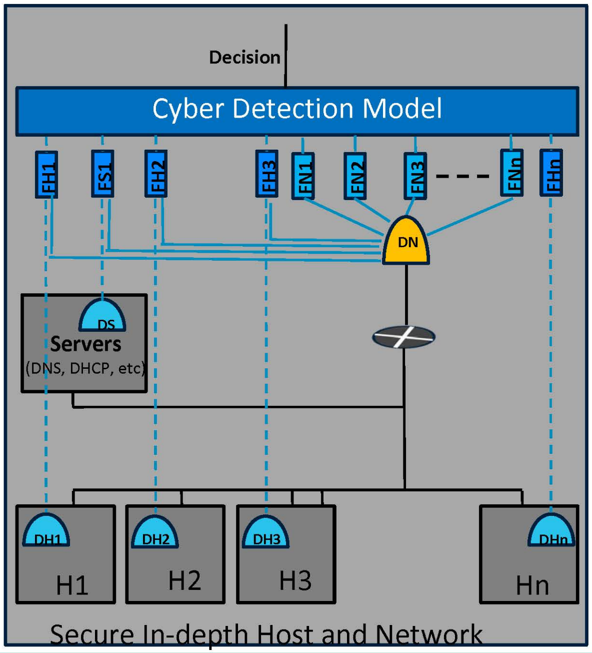
Adversarial setting:

- Think adversary disabling anti-virus program upon host infection
- Adversary will undermine the input (and possibly the output) of the ML model
  - Disable sensors/data upon initial intrusion
  - Disable feature input to the ML model
  - *Fake* feature value
- Trigger negative detection decision (false negative)

Conundrum:

- Perimeter ML IDS is *secure* but *imperceptible*
- In-depth ML IDS *perceptible* but *unsecure*

Approach: Network Based Intrusion Detection over Opaque Data

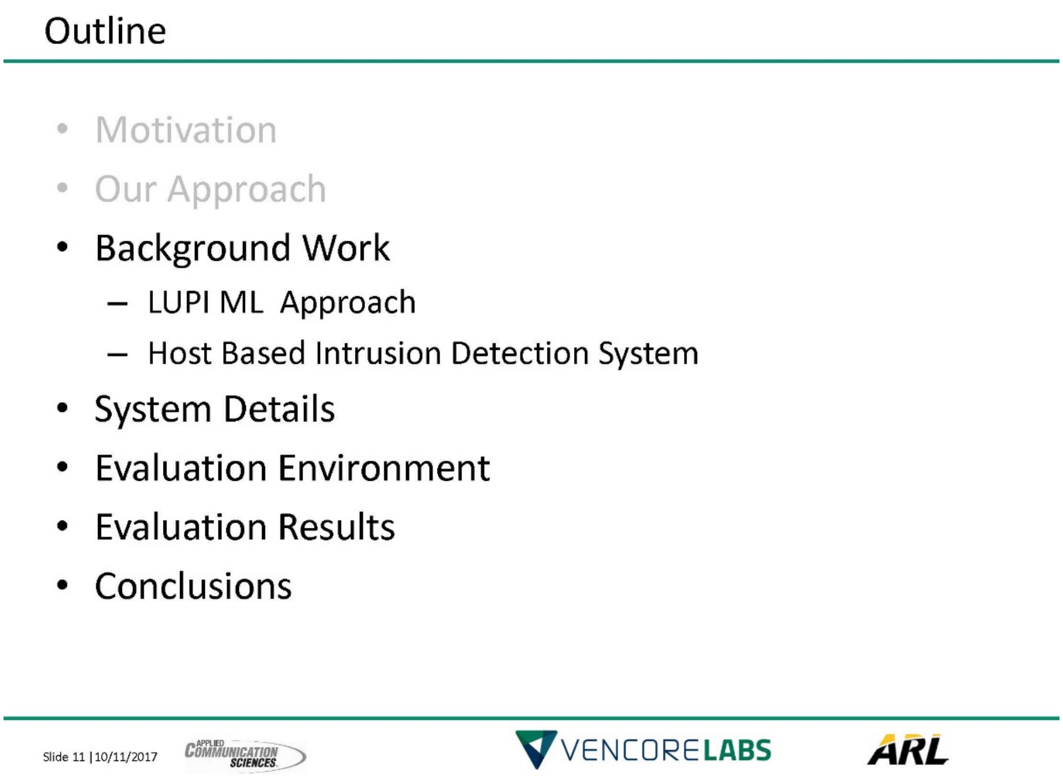


In-depth ML:

- Perceptive to host state, server state, network state
- Highly accurate model of System under Defense

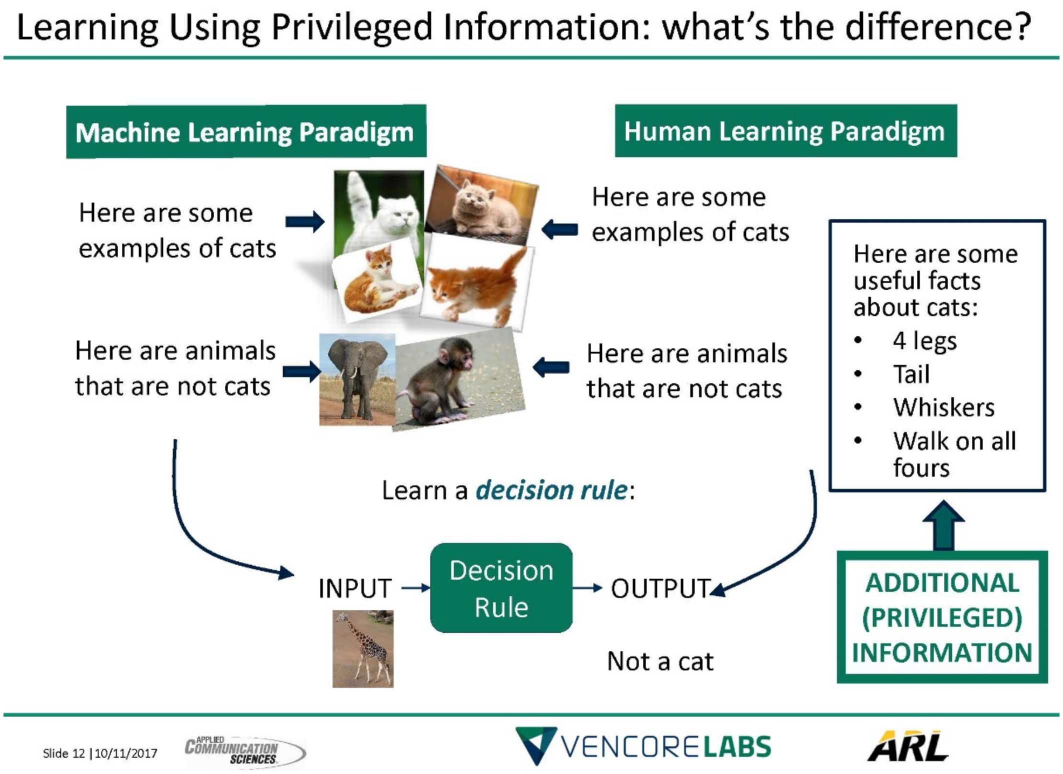
Secure ML:

- Vulnerable features not measured directly at runtime
- Estimated instead at secure location from common data source



Outline

- Motivation
- Our Approach
- Background Work
  - LUPI ML Approach
  - Host Based Intrusion Detection System
- System Details
- Evaluation Environment
- Evaluation Results
- Conclusions



Learning Using Privileged Information: what's the difference?

Machine Learning Paradigm

Here are some examples of cats

Here are animals that are not cats

Human Learning Paradigm

Here are some examples of cats

Here are animals that are not cats

Here are some useful facts about cats:

- 4 legs
- Tail
- Whiskers
- Walk on all fours

ADDITIONAL (PRIVILEGED) INFORMATION

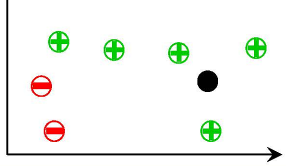
Learn a ***decision rule***:

INPUT → Decision Rule → OUTPUT

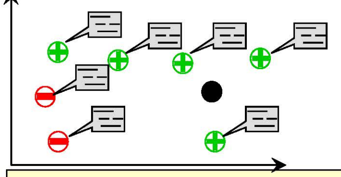
Not a cat

Learning Using Privileged Information (LUPI)

CLASSICAL MACHINE LEARNING



LEARNING USING PRIVILEGED INFORMATION



- Given training data (observations, facts)

  $(x_1, y_1), \ldots, (x_L, y_L)$ ⊕ ⊖

- Generalize data to a rule (function)

  where $y = f(x)$

  and $x \in X$ $y \in \{-1,+1\}$ ●

- Given training data (observations, facts)

  $(x_1, y_1), \ldots, (x_L, y_L)$ ⊕ ⊖

  and additional privileged data

  $x_1^*, \ldots x_L^*$

- Generalize data to a rule (function)

  $y = f(x)$ ●

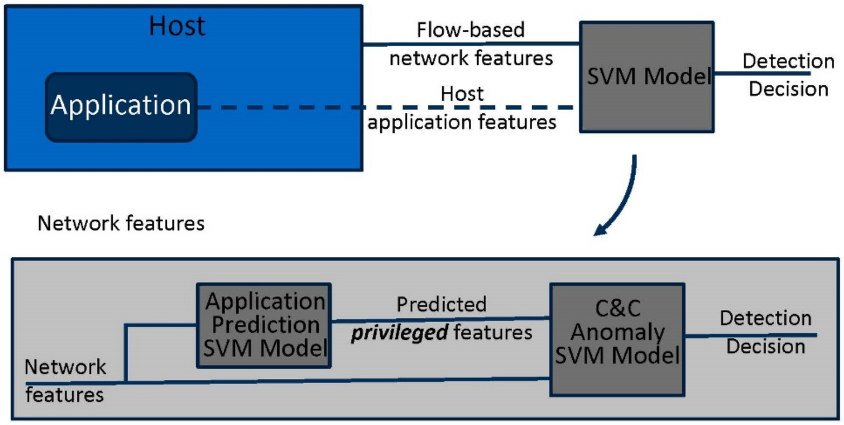
  where $x \in X, x^* \in X^*$ and $y \in \{-1,+1\}$

- Classical machine learning: training data and test data are from the same space, with have same attributes etc.
- Solved by SVM algorithm

- New paradigm of learning with privileged information: additional information is available ONLY with training data, but NOT with test data
- Solved by SVM+ algorithm

Previous Work: Host Based Intrusion Detection *

- Detect botnet C&C communication (FastFlux, DGA) using host-based network tap



Host

Application

Flow-based network features

Host application features

SVM Model

Detection Decision

Network features

Application Prediction SVM Model

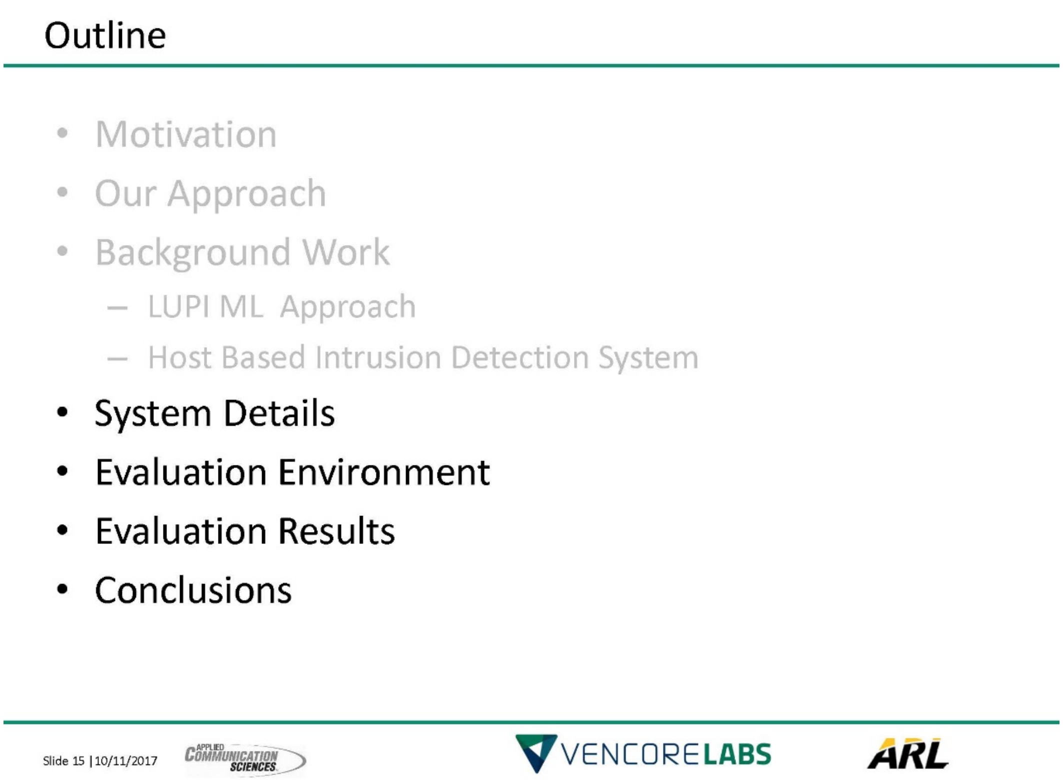
Predicted *privileged* features

C&C Anomaly SVM Model

Network features

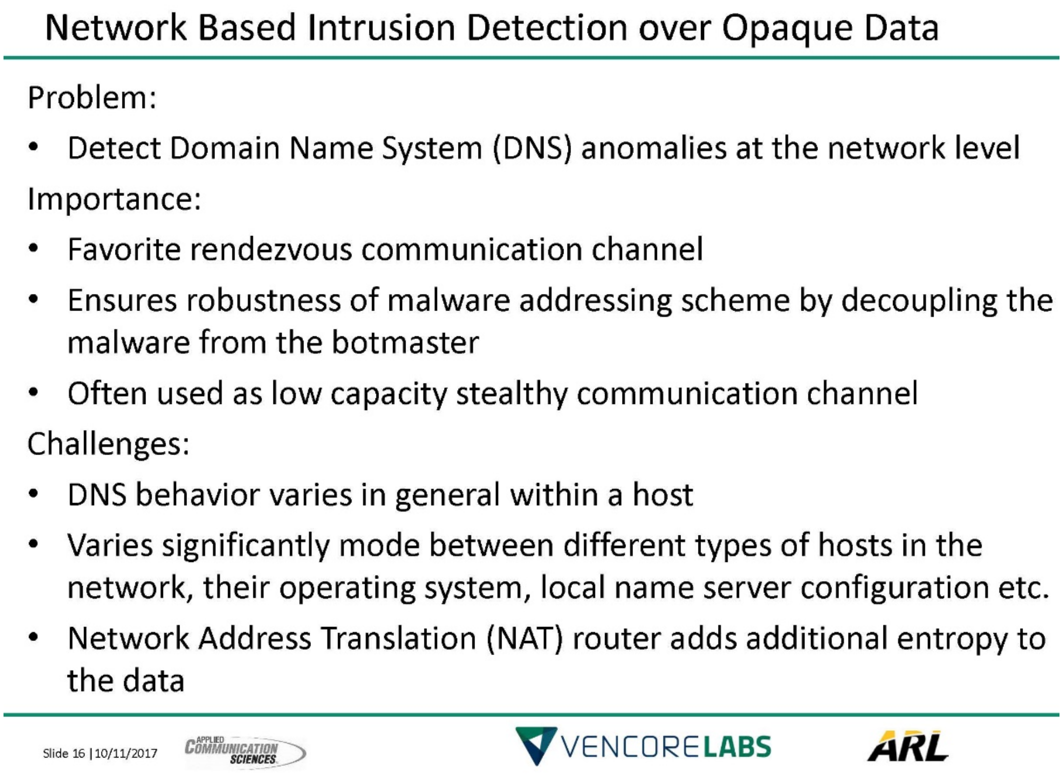
Detection Decision

*A.Sapello, C. Serban, R. Chadha, and R. Izmailov, "Application of Learning Using Privileged Information(LUPI): Botnet Detection", in the Workshop on Network Security Analytics and Automation (NSAA), 2017



Outline

- Motivation
- Our Approach
- Background Work
  - LUPI ML Approach
  - Host Based Intrusion Detection System
- System Details
- Evaluation Environment
- Evaluation Results
- Conclusions



Network Based Intrusion Detection over Opaque Data

Problem:

- Detect Domain Name System (DNS) anomalies at the network level

Importance:

- Favorite rendezvous communication channel
- Ensures robustness of malware addressing scheme by decoupling the malware from the botmaster
- Often used as low capacity stealthy communication channel

Challenges:

- DNS behavior varies in general within a host
- Varies significantly mode between different types of hosts in the network, their operating system, local name server configuration etc.
- Network Address Translation (NAT) router adds additional entropy to the data

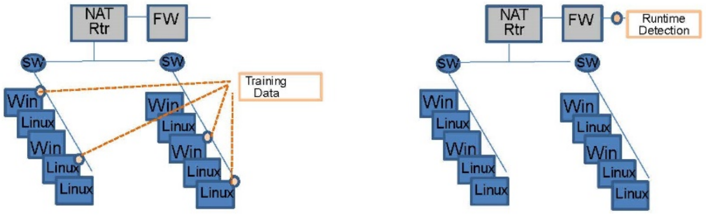
System Details

Basic features:

- 88 IP features based on inter-packet arrival time and packet size centrality measures (i.e., minimum, maximum, average, and standard deviation)
- Split into forward and backward packet metrics
- Combined across related flows into a flow family (session).
- 6 DNS features (# of DNS requests, #DNS refused-NX, #flows)

Privileged Features:

- Type of Operating System of the host issuing a flow family.



(a) Training within Enterprise

(b) Detection outside of enterprise

System Details (cont'ed)

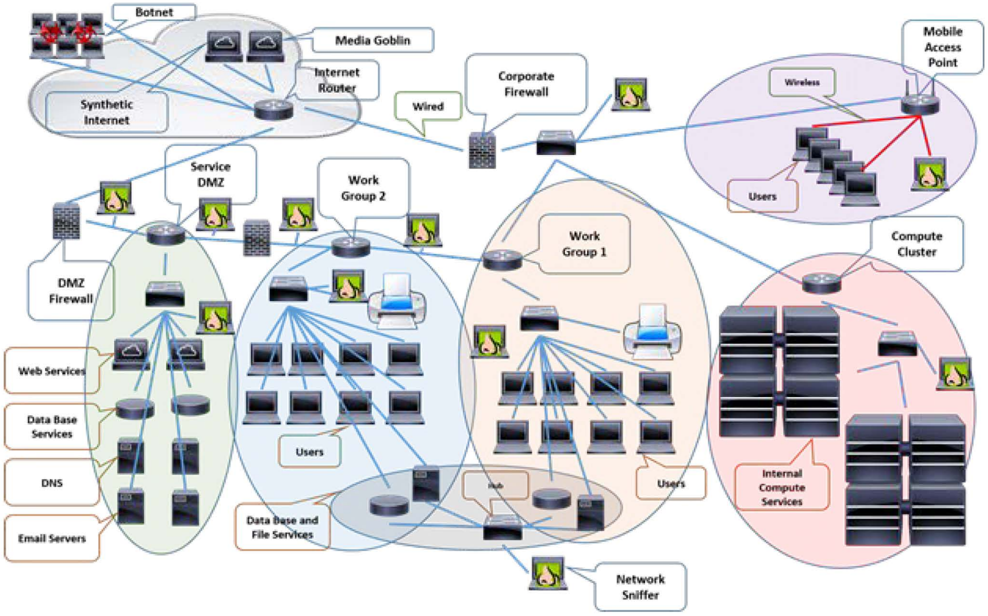
Privileged feature: flow family OS

- Highly informative of the expected DNS behavior of the communication stack, and for the problem at hand:
  - Insight: different OS's have different DNS caching policies (i.e. how long the stack remembers the resolution between a name and an IP address)
  - Dominates the number and types of requests visible upstream
- Unavailable at runtime for secure perimeter detection:
  - NAT routers obscure the source address of IP packets hence no static OS-to-IP mapping
  - Further, HTTPS and other protocols may obscure the agent/type/OS of the requester
- It can be inferred, however, *via training*:
  - From same flow-based network features
  - And from *privileged* labels collected at training time

Training Details

- Support Vector Machines (SVM)
- Privileged features predicted via Two-Class SVM classification trained on data collected on host link (with Win7/LinuxFC20 OS as label)
- DNS anomaly predicted via One-Class SVM classification trained on data collected on the NAT up-link
- Training and evaluation data covering each over 10+ hours of flows from 25+ hosts
- Performed in CyberVAN in a Corporate Environment Scenario:
  - Synthetic environment
  - See next

Training and Evaluation Scenario

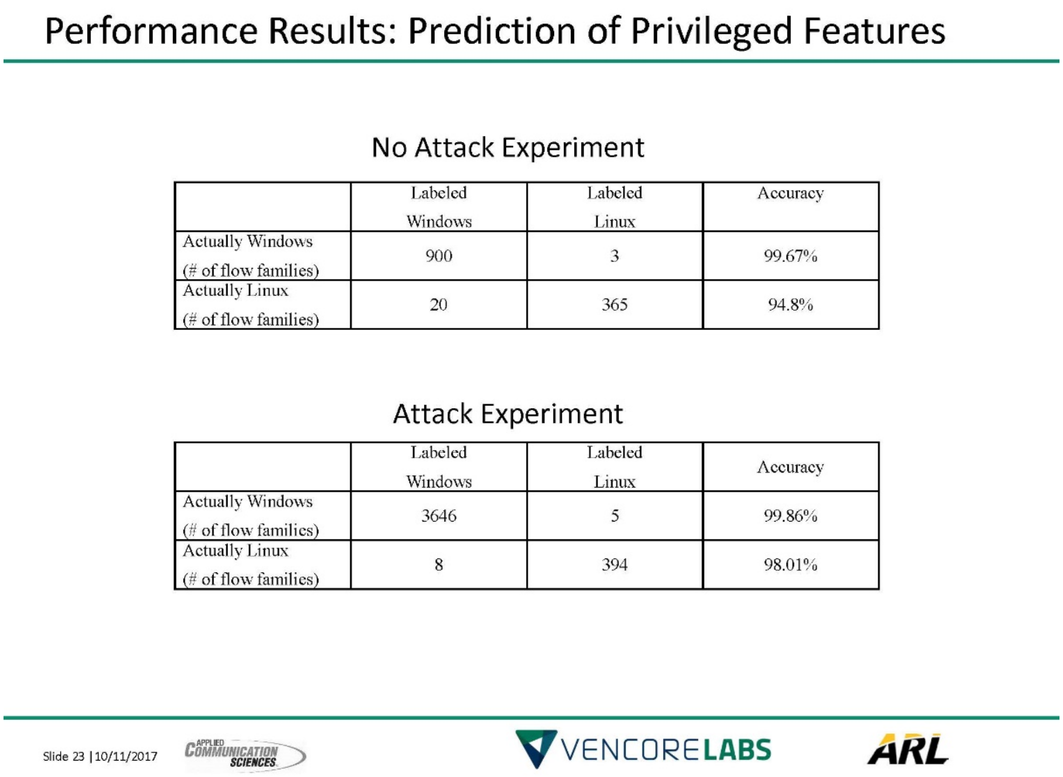


Adversary Modeling

- Adversary creates spurious DNS queries
- Queries *synchronized* with existing data flow
- Malware wakes up randomly:
  - If background IP traffic, send one spurious DNS query with given probability
  - Otherwise go to sleep
- Models DGA type communication but with added covert measures

Outline

- Motivation
- Our Approach
- Background Work
  - LUPI ML Approach
  - Host Based Intrusion Detection System
- System Details
- Evaluation Environment
- Evaluation Results
- Conclusions



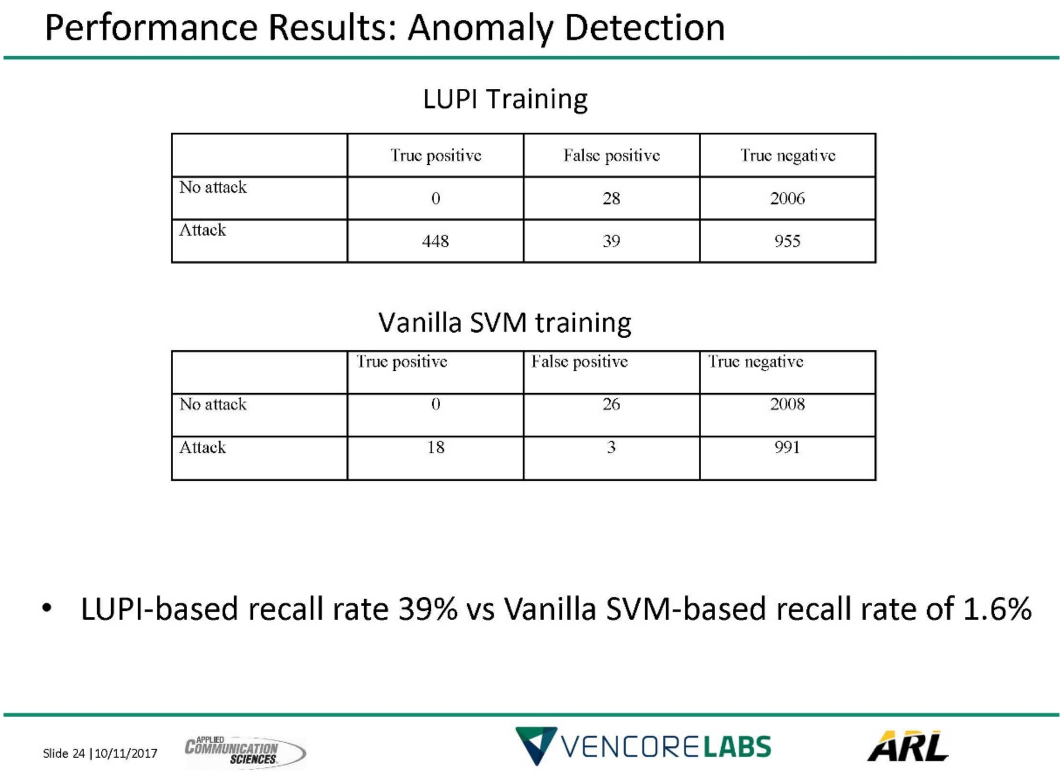
Performance Results: Prediction of Privileged Features

No Attack Experiment

	Labeled Windows	Labeled Linux	Accuracy
Actually Windows (# of flow families)	900	3	99.67%
Actually Linux (# of flow families)	20	365	94.8%

Attack Experiment

	Labeled Windows	Labeled Linux	Accuracy
Actually Windows (# of flow families)	3646	5	99.86%
Actually Linux (# of flow families)	8	394	98.01%



Performance Results: Anomaly Detection

LUPI Training

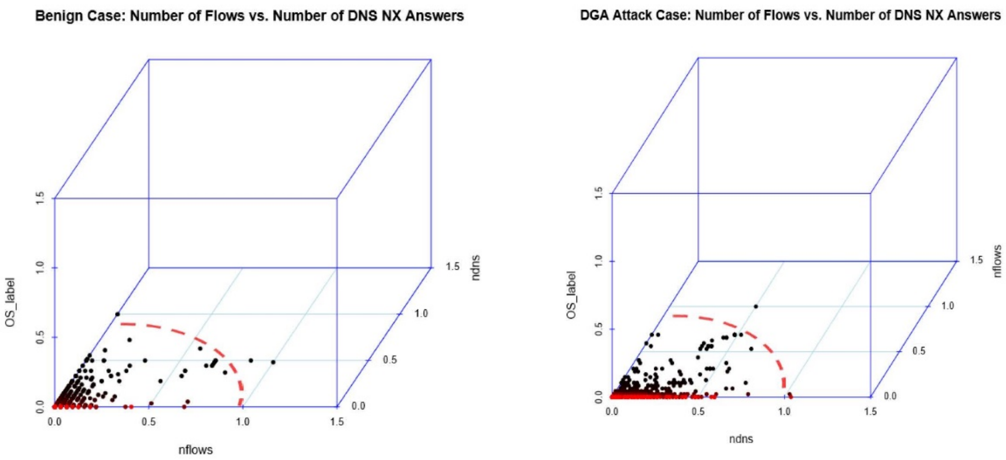
	True positive	False positive	True negative
No attack	0	28	2006
Attack	448	39	955

Vanilla SVM training

	True positive	False positive	True negative
No attack	0	26	2008
Attack	18	3	991

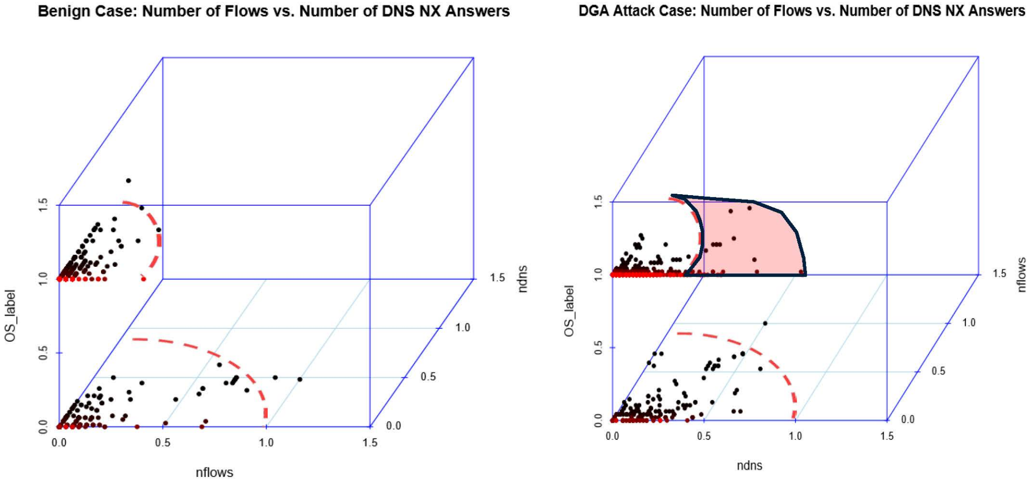
- LUPI-based recall rate 39% vs Vanilla SVM-based recall rate of 1.6%

Geometrical Interpretation


Benign Case: Number of Flows vs. Number of DNS NX Answers

DGA Attack Case: Number of Flows vs. Number of DNS NX Answers

Geometrical Interpretation (cont'ed)

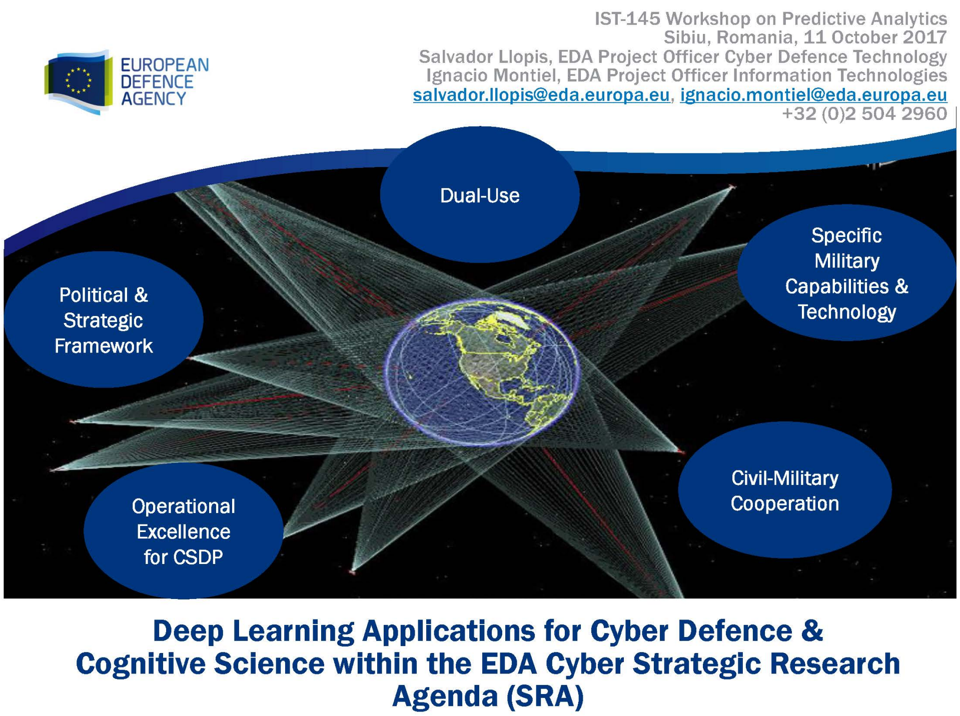

Benign Case: Number of Flows vs. Number of DNS NX Answers

DGA Attack Case: Number of Flows vs. Number of DNS NX Answers

Conclusions

- New approach for improving the accuracy of ML models via privileged features available only during training
  - Capable of detecting ATP-type stealthy malicious behavior
- Provides safety against feature tampering by adversary
  - Operates on perimeter in upstream safe network
- DNS Anomaly model implementation:
  - Detects stealthy DNS spurious requests
  - Operates in the up-link of the NAT router on opaque data
- Future work:
  - Improved flow family grouping
  - Improved DNS-to-flow imputation
  - Additional privileged features

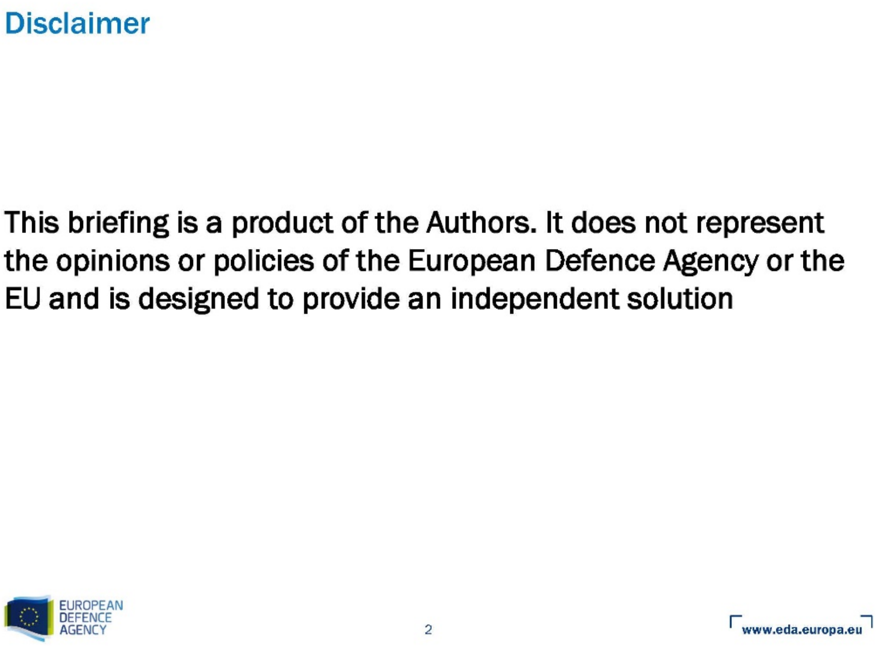
Thank You !



IST-145 Workshop on Predictive Analytics
Sibiu, Romania, 11 October 2017
Salvador Llopis, EDA Project Officer Cyber Defence Technology
Ignacio Montiel, EDA Project Officer Information Technologies
salvador.llopis@eda.europa.eu, ignacio.montiel@eda.europa.eu
+32 (0)2 504 2960

Dual-Use
Specific Military Capabilities & Technology
Political & Strategic Framework
Operational Excellence for CSDP
Civil-Military Cooperation

Deep Learning Applications for Cyber Defence & Cognitive Science within the EDA Cyber Strategic Research Agenda (SRA)

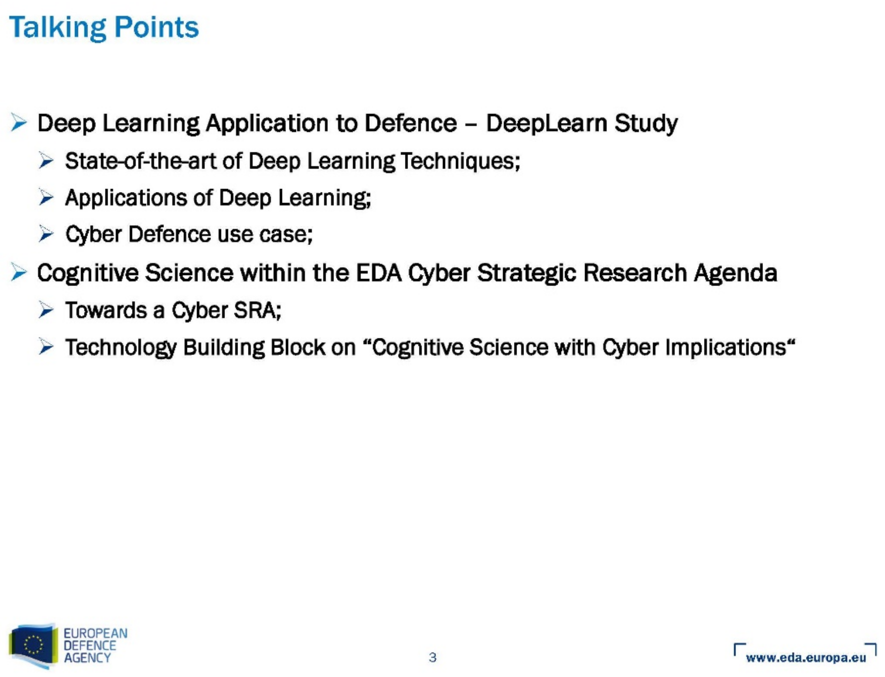


Disclaimer

This briefing is a product of the Authors. It does not represent the opinions or policies of the European Defence Agency or the EU and is designed to provide an independent solution

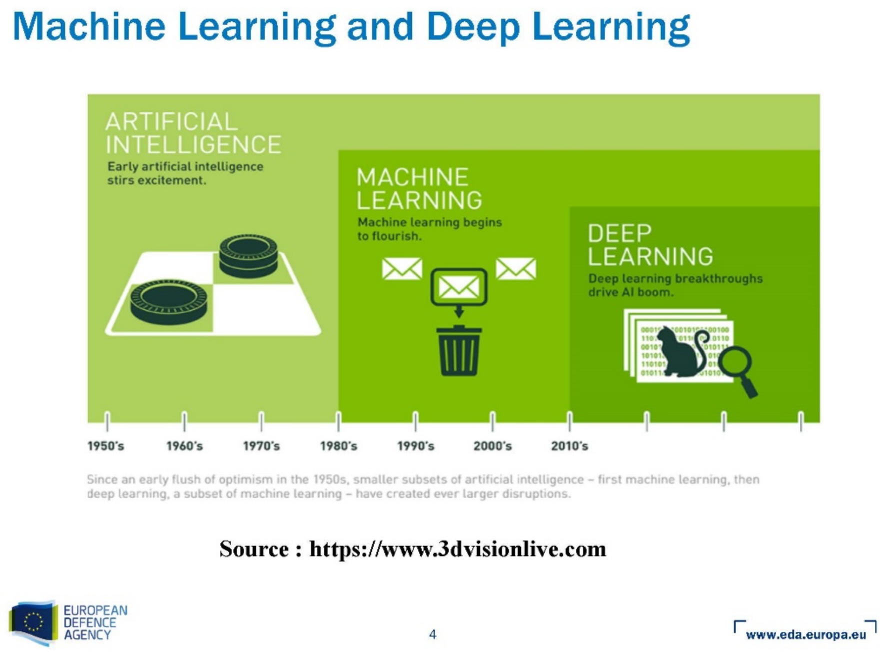
2

www.eda.europa.eu



Talking Points

- Deep Learning Application to Defence – DeepLearn Study
  - State-of-the-art of Deep Learning Techniques;
  - Applications of Deep Learning;
  - Cyber Defence use case;
- Cognitive Science within the EDA Cyber Strategic Research Agenda
  - Towards a Cyber SRA;
  - Technology Building Block on “Cognitive Science with Cyber Implications“

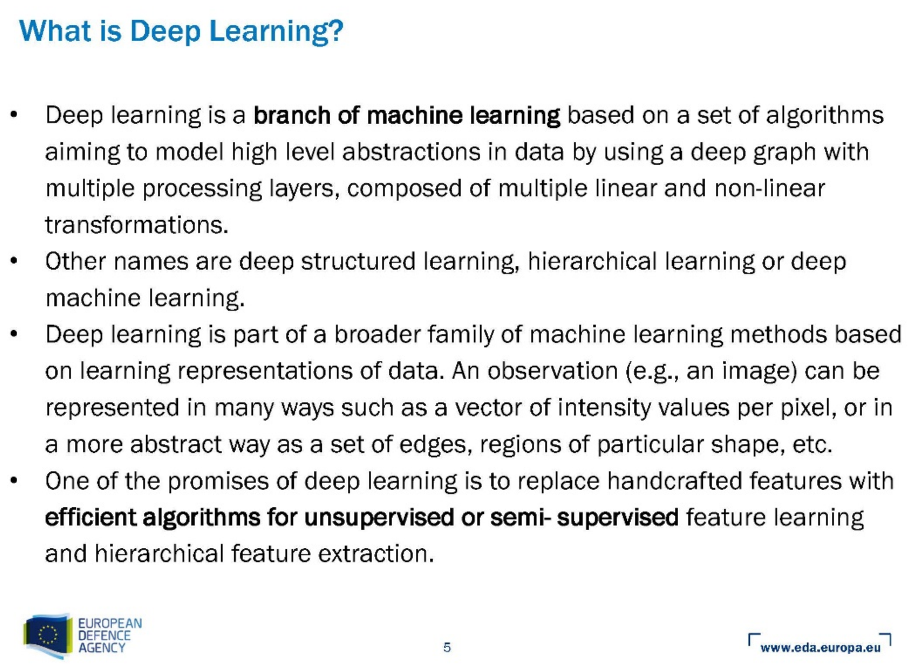


Machine Learning and Deep Learning

ARTIFICIAL INTELLIGENCE
Early artificial intelligence stirs excitement.
MACHINE LEARNING
Machine learning begins to flourish.
DEEP LEARNING
Deep learning breakthroughs drive AI boom.
1950's
1960's
1970's
1980's
1990's
2000's
2010's

Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

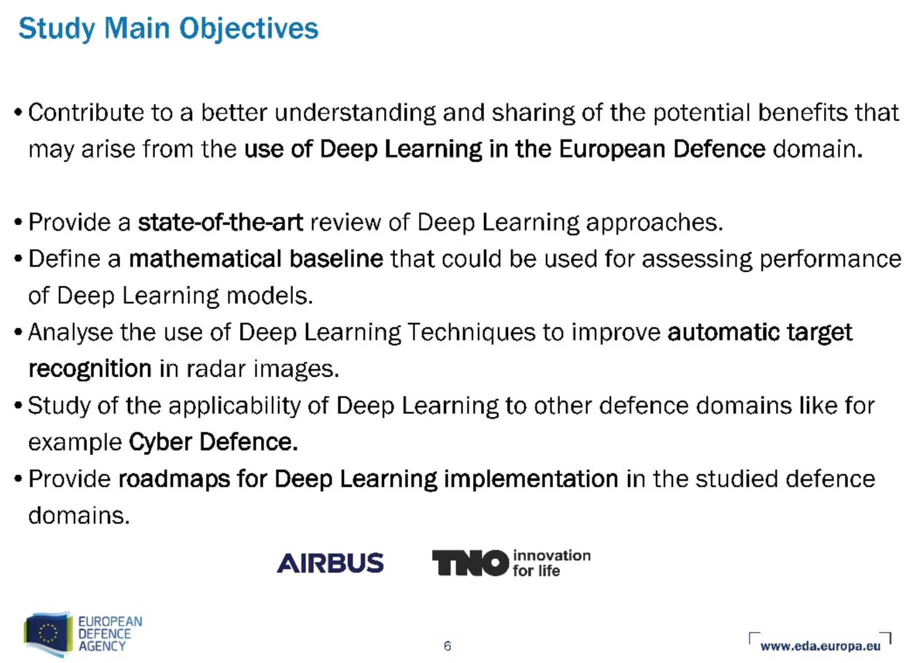
Source : https://www.3dvisionlive.com



What is Deep Learning?

- Deep learning is a **branch of machine learning** based on a set of algorithms aiming to model high level abstractions in data by using a deep graph with multiple processing layers, composed of multiple linear and non-linear transformations.
- Other names are deep structured learning, hierarchical learning or deep machine learning.
- Deep learning is part of a broader family of machine learning methods based on learning representations of data. An observation (e.g., an image) can be represented in many ways such as a vector of intensity values per pixel, or in a more abstract way as a set of edges, regions of particular shape, etc.
- One of the promises of deep learning is to replace handcrafted features with **efficient algorithms for unsupervised or semi- supervised** feature learning and hierarchical feature extraction.

EUROPEAN DEFENCE AGENCY 5 www.eda.europa.eu



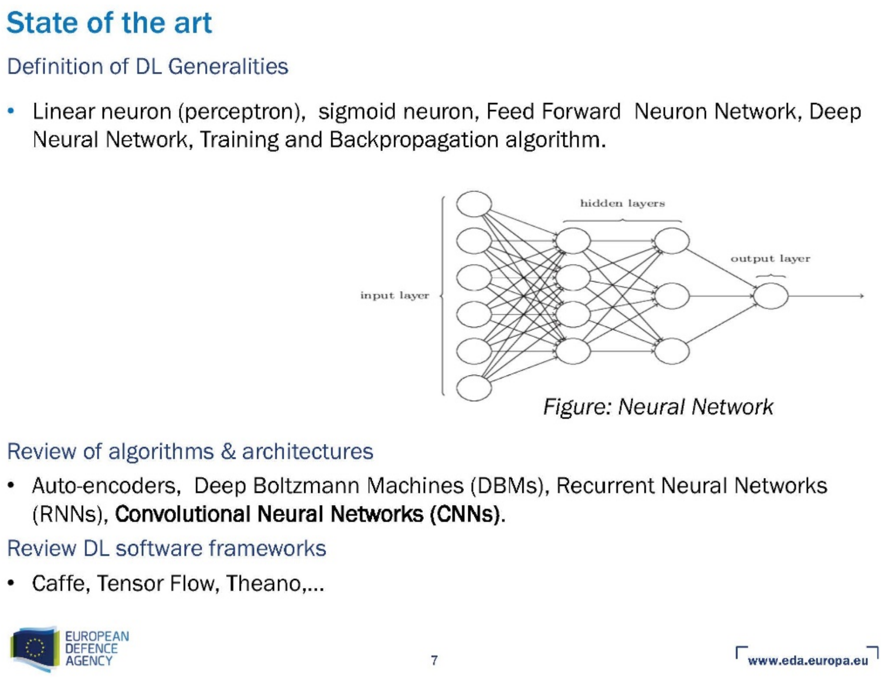
Study Main Objectives

- Contribute to a better understanding and sharing of the potential benefits that may arise from the **use of Deep Learning in the European Defence** domain.

- Provide a **state-of-the-art** review of Deep Learning approaches.
- Define a **mathematical baseline** that could be used for assessing performance of Deep Learning models.
- Analyse the use of Deep Learning Techniques to improve **automatic target recognition** in radar images.
- Study of the applicability of Deep Learning to other defence domains like for example **Cyber Defence.**
- Provide **roadmaps for Deep Learning implementation** in the studied defence domains.

AIRBUS TNO innovation for life

EUROPEAN DEFENCE AGENCY 6 www.eda.europa.eu



State of the art

Definition of DL Generalities

- Linear neuron (perceptron), sigmoid neuron, Feed Forward Neuron Network, Deep Neural Network, Training and Backpropagation algorithm.

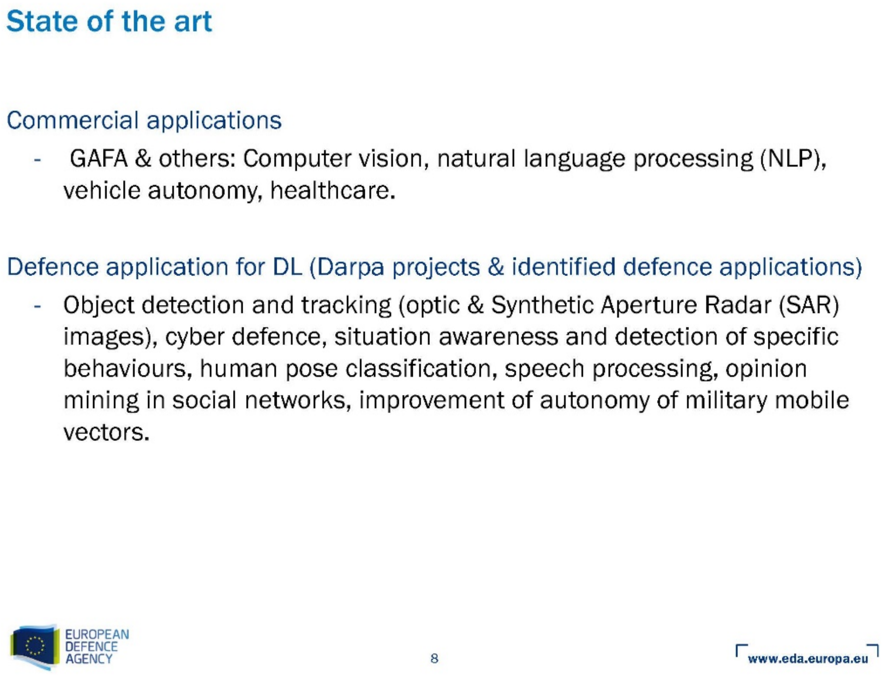
Figure: Neural Network

Review of algorithms & architectures

- Auto-encoders, Deep Boltzmann Machines (DBMs), Recurrent Neural Networks (RNNs), **Convolutional Neural Networks (CNNs)**.

Review DL software frameworks

- Caffe, Tensor Flow, Theano,...



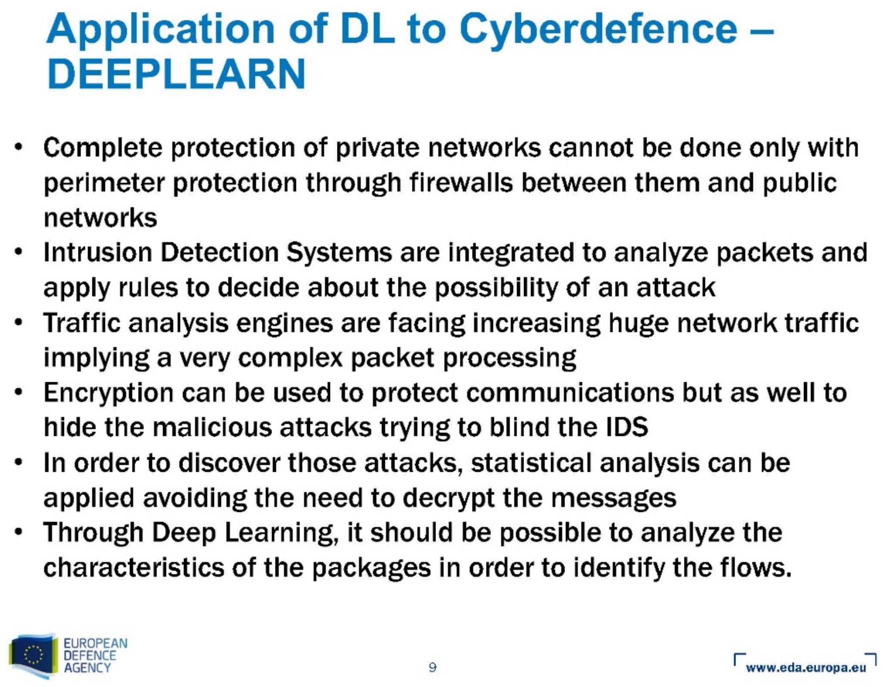
State of the art

Commercial applications

- GAFA & others: Computer vision, natural language processing (NLP), vehicle autonomy, healthcare.

Defence application for DL (Darpa projects & identified defence applications)

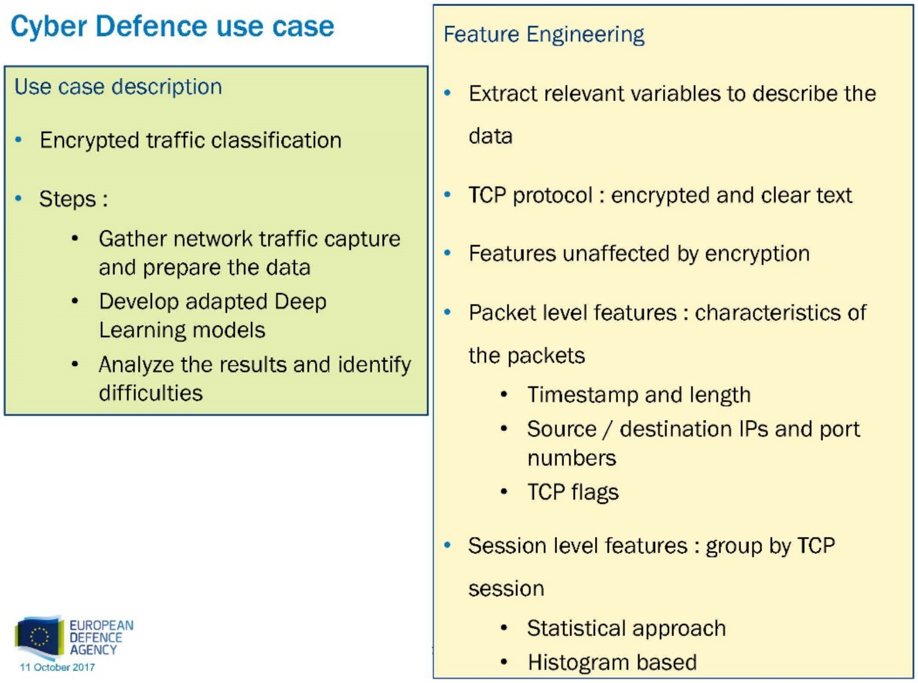
- Object detection and tracking (optic & Synthetic Aperture Radar (SAR) images), cyber defence, situation awareness and detection of specific behaviours, human pose classification, speech processing, opinion mining in social networks, improvement of autonomy of military mobile vectors.



Application of DL to Cyberdefence – DEEPLEARN

- Complete protection of private networks cannot be done only with perimeter protection through firewalls between them and public networks
- Intrusion Detection Systems are integrated to analyze packets and apply rules to decide about the possibility of an attack
- Traffic analysis engines are facing increasing huge network traffic implying a very complex packet processing
- Encryption can be used to protect communications but as well to hide the malicious attacks trying to blind the IDS
- In order to discover those attacks, statistical analysis can be applied avoiding the need to decrypt the messages
- Through Deep Learning, it should be possible to analyze the characteristics of the packages in order to identify the flows.

EUROPEAN DEFENCE AGENCY 9 www.eda.europa.eu



Cyber Defence use case

Use case description

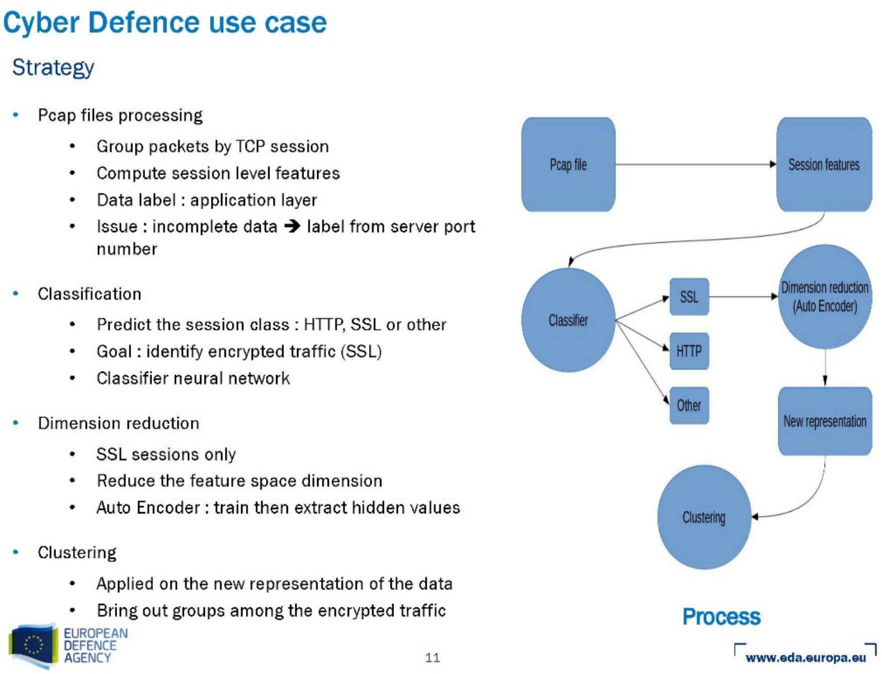
- Encrypted traffic classification
- Steps :
  - Gather network traffic capture and prepare the data
  - Develop adapted Deep Learning models
  - Analyze the results and identify difficulties

Feature Engineering

- Extract relevant variables to describe the data
- TCP protocol : encrypted and clear text
- Features unaffected by encryption
- Packet level features : characteristics of the packets
  - Timestamp and length
  - Source / destination IPs and port numbers
  - TCP flags
- Session level features : group by TCP session
  - Statistical approach
  - Histogram based

EUROPEAN DEFENCE AGENCY

11 October 2017

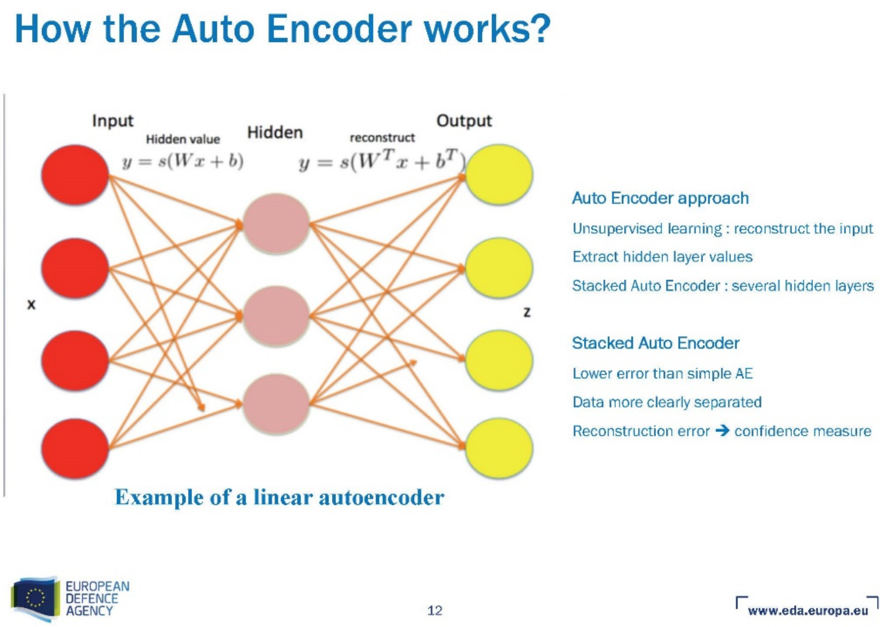


Cyber Defence use case

Strategy

- Pcap files processing
  - Group packets by TCP session
  - Compute session level features
  - Data label : application layer
  - Issue : incomplete data ➔ label from server port number
- Classification
  - Predict the session class : HTTP, SSL or other
  - Goal : identify encrypted traffic (SSL)
  - Classifier neural network
- Dimension reduction
  - SSL sessions only
  - Reduce the feature space dimension
  - Auto Encoder : train then extract hidden values
- Clustering
  - Applied on the new representation of the data
  - Bring out groups among the encrypted traffic

Process



How the Auto Encoder works?

Example of a linear autoencoder

Auto Encoder approach

Unsupervised learning : reconstruct the input

Extract hidden layer values

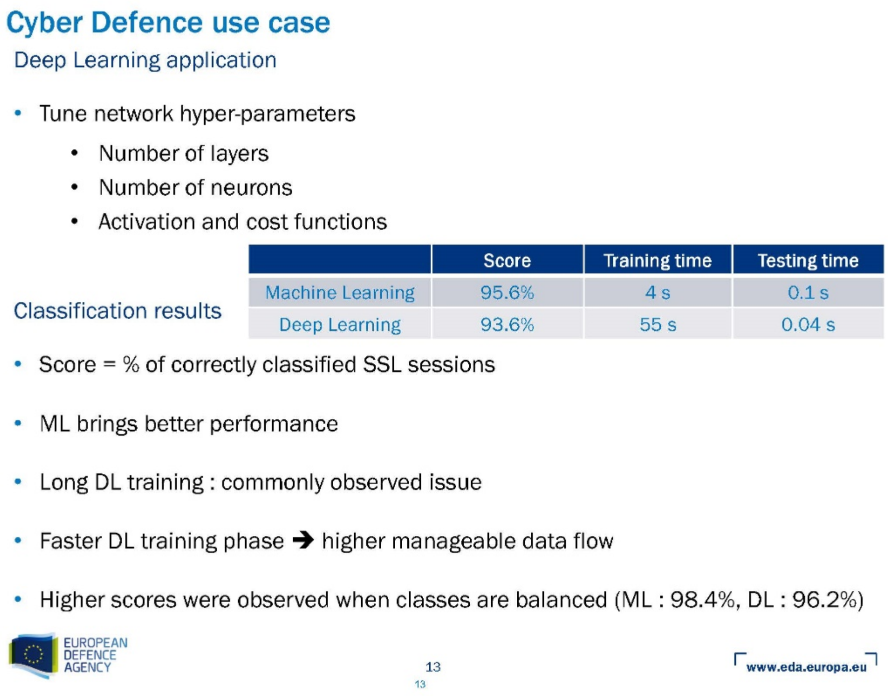
Stacked Auto Encoder : several hidden layers

Stacked Auto Encoder

Lower error than simple AE

Data more clearly separated

Reconstruction error ➔ confidence measure



Cyber Defence use case

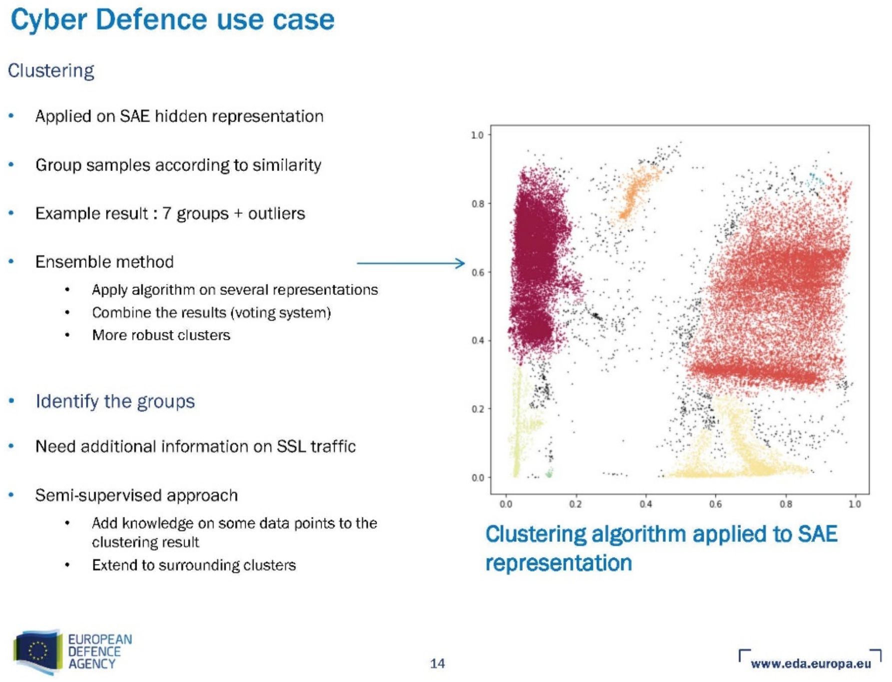
Deep Learning application

- Tune network hyper-parameters
  - Number of layers
  - Number of neurons
  - Activation and cost functions

Classification results

	Score	Training time	Testing time
Machine Learning	95.6%	4 s	0.1 s
Deep Learning	93.6%	55 s	0.04 s

- Score = % of correctly classified SSL sessions
- ML brings better performance
- Long DL training : commonly observed issue
- Faster DL training phase ➔ higher manageable data flow
- Higher scores were observed when classes are balanced (ML : 98.4%, DL : 96.2%)

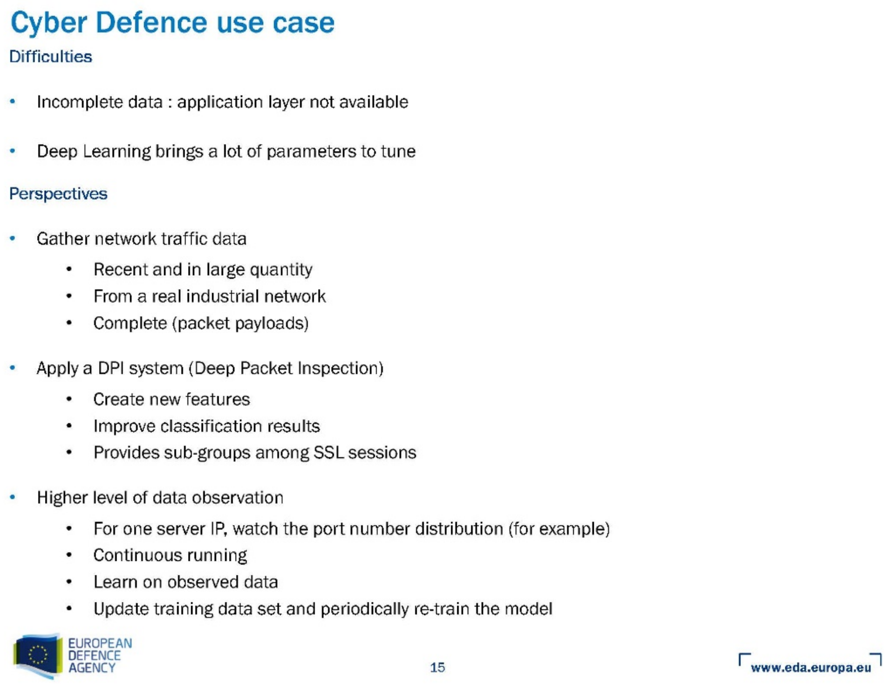


Cyber Defence use case

Clustering

- Applied on SAE hidden representation
- Group samples according to similarity
- Example result : 7 groups + outliers
- Ensemble method
  - Apply algorithm on several representations
  - Combine the results (voting system)
  - More robust clusters
- Identify the groups
- Need additional information on SSL traffic
- Semi-supervised approach
  - Add knowledge on some data points to the clustering result
  - Extend to surrounding clusters

Clustering algorithm applied to SAE representation



Cyber Defence use case

Difficulties

- Incomplete data : application layer not available
- Deep Learning brings a lot of parameters to tune

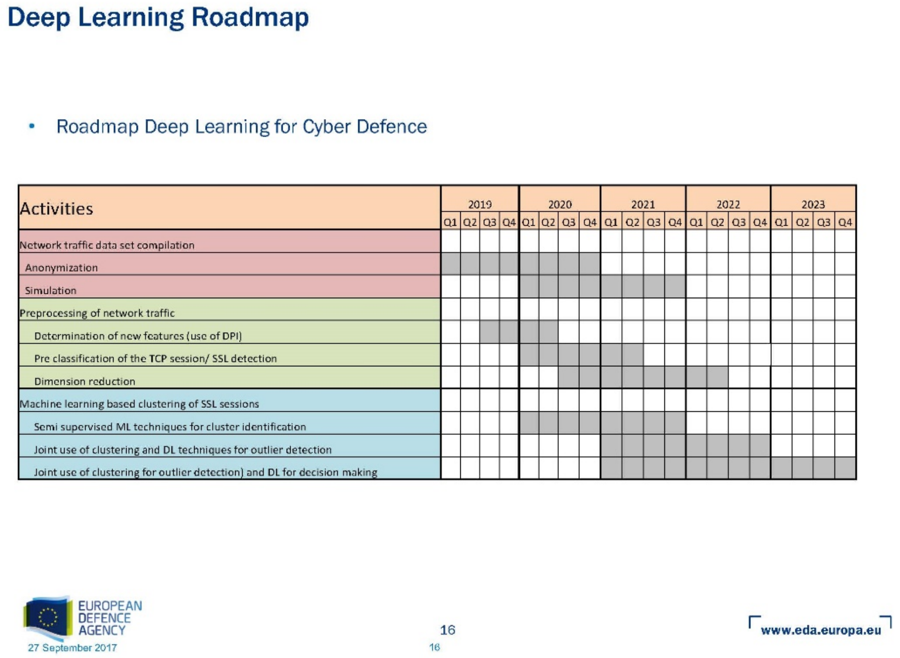
Perspectives

- Gather network traffic data
  - Recent and in large quantity
  - From a real industrial network
  - Complete (packet payloads)
- Apply a DPI system (Deep Packet Inspection)
  - Create new features
  - Improve classification results
  - Provides sub-groups among SSL sessions
- Higher level of data observation
  - For one server IP, watch the port number distribution (for example)
  - Continuous running
  - Learn on observed data
  - Update training data set and periodically re-train the model

EUROPEAN DEFENCE AGENCY

15

www.eda.europa.eu



Deep Learning Roadmap

- Roadmap Deep Learning for Cyber Defence

Activities	2019				2020				2021				2022				2023			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Network traffic data set compilation																				
Anonymization	▨	▨	▨	▨	▨	▨	▨	▨												
Simulation					▨	▨	▨	▨	▨	▨	▨	▨								
Preprocessing of network traffic																				
Determination of new features (use of DPI)			▨	▨	▨	▨														
Pre classification of the TCP session/ SSL detection					▨	▨	▨	▨	▨	▨										
Dimension reduction							▨	▨	▨	▨	▨	▨	▨	▨						
Machine learning based clustering of SSL sessions																				
Semi supervised ML techniques for cluster identification					▨	▨	▨	▨	▨	▨	▨	▨								
Joint use of clustering and DL techniques for outlier detection									▨	▨	▨	▨	▨	▨	▨	▨				
Joint use of clustering for outlier detection) and DL for decision making									▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨

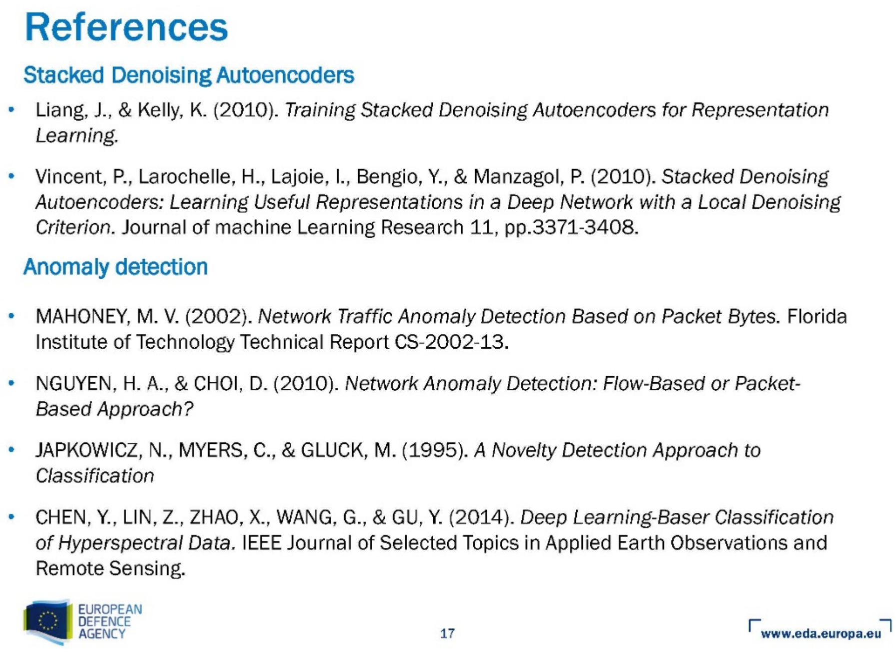
EUROPEAN DEFENCE AGENCY

27 September 2017

16

16

www.eda.europa.eu



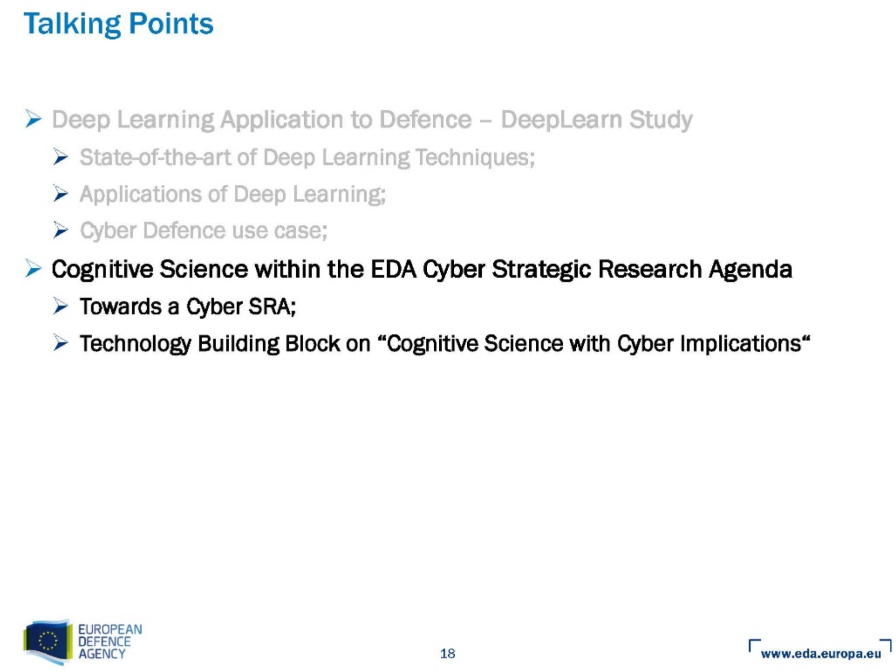
References

Stacked Denoising Autoencoders

- Liang, J., & Kelly, K. (2010). *Training Stacked Denoising Autoencoders for Representation Learning.*
- Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. (2010). *Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion.* Journal of machine Learning Research 11, pp.3371-3408.

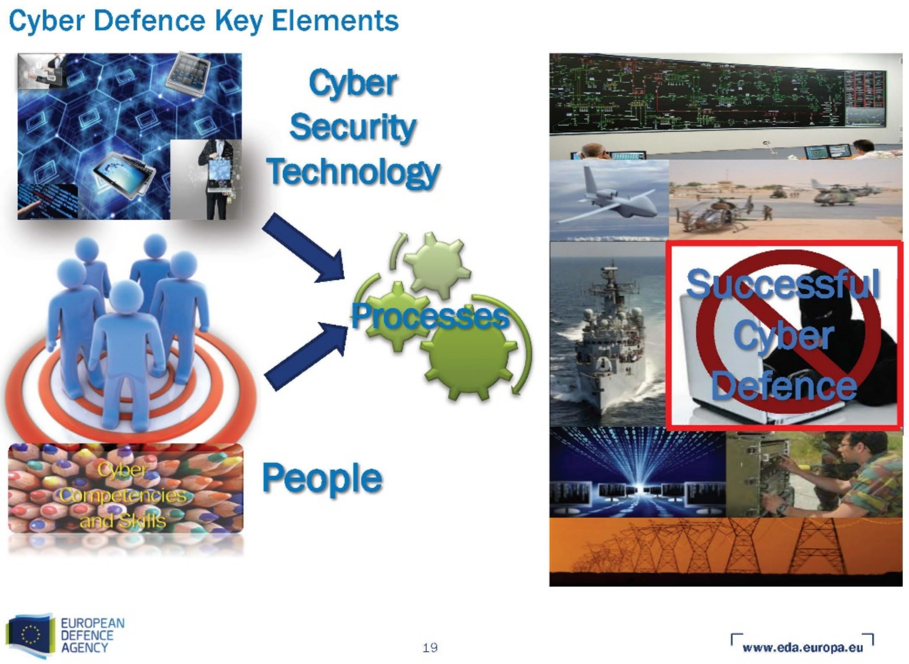
Anomaly detection

- MAHONEY, M. V. (2002). *Network Traffic Anomaly Detection Based on Packet Bytes.* Florida Institute of Technology Technical Report CS-2002-13.
- NGUYEN, H. A., & CHOI, D. (2010). *Network Anomaly Detection: Flow-Based or Packet-Based Approach?*
- JAPKOWICZ, N., MYERS, C., & GLUCK, M. (1995). *A Novelty Detection Approach to Classification*
- CHEN, Y., LIN, Z., ZHAO, X., WANG, G., & GU, Y. (2014). *Deep Learning-Baser Classification of Hyperspectral Data.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.



Talking Points

- ➢ Deep Learning Application to Defence – DeepLearn Study
  - ➢ State-of-the-art of Deep Learning Techniques;
  - ➢ Applications of Deep Learning;
  - ➢ Cyber Defence use case;
- ➢ Cognitive Science within the EDA Cyber Strategic Research Agenda
  - ➢ Towards a Cyber SRA;
  - ➢ Technology Building Block on “Cognitive Science with Cyber Implications“



Cyber Defence Key Elements

Cyber Security Technology

Processes

People

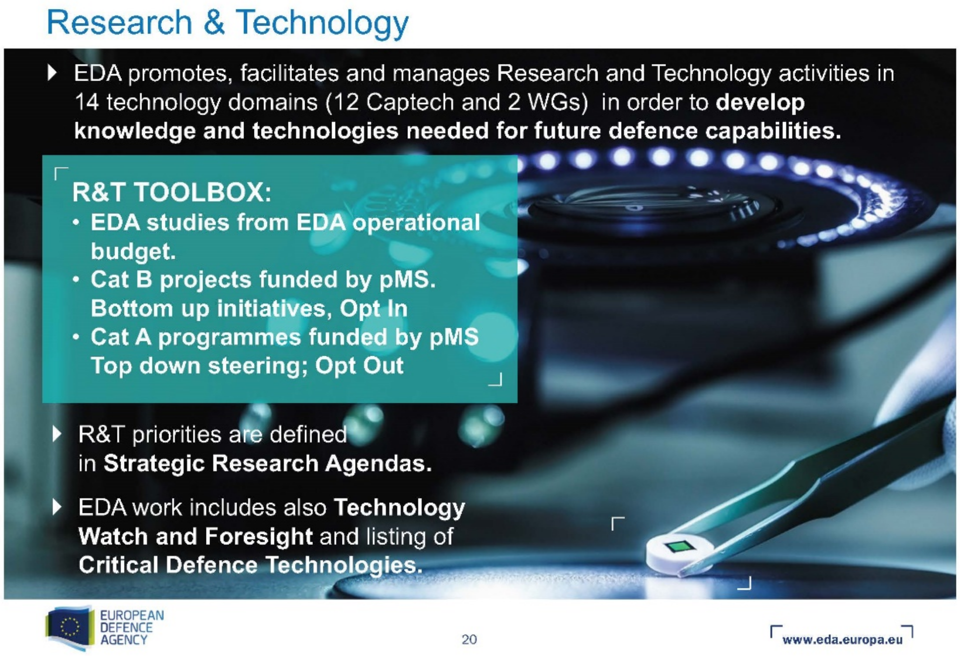
Cyber Competencies and Skills

Successful Cyber Defence

EUROPEAN DEFENCE AGENCY

19

www.eda.europa.eu



Research & Technology

- EDA promotes, facilitates and manages Research and Technology activities in 14 technology domains (12 Captech and 2 WGs) in order to **develop knowledge and technologies needed for future defence capabilities.**

R&T TOOLBOX:

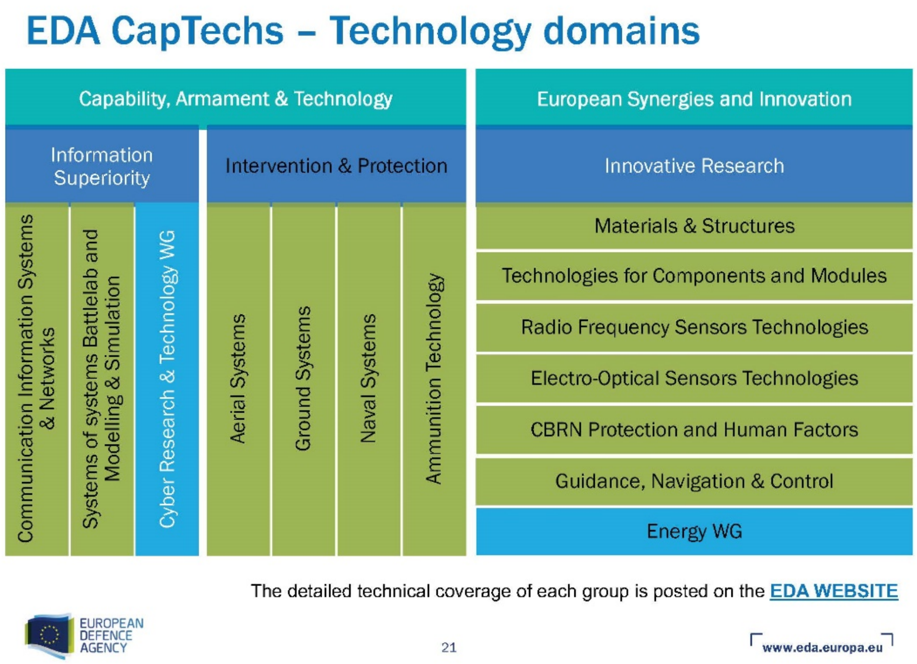
- **EDA studies from EDA operational budget.**
- **Cat B projects funded by pMS. Bottom up initiatives, Opt In**
- **Cat A programmes funded by pMS Top down steering; Opt Out**

- R&T priorities are defined in **Strategic Research Agendas.**
- EDA work includes also **Technology Watch and Foresight** and listing of **Critical Defence Technologies.**

EUROPEAN DEFENCE AGENCY

20

www.eda.europa.eu



EDA CapTechs – Technology domains

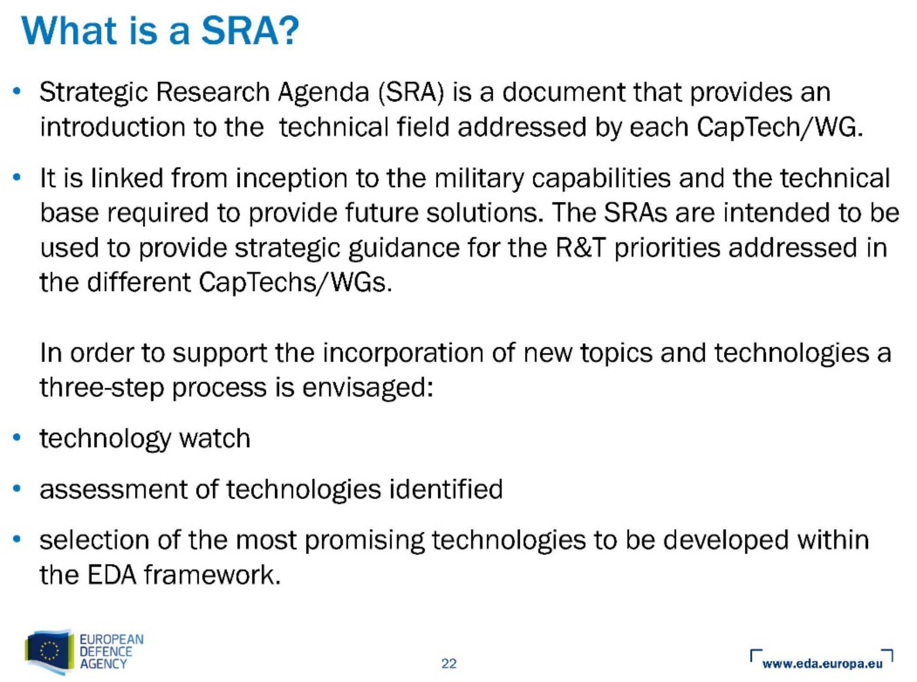
Capability, Armament & Technology		European Synergies and Innovation
Information Superiority	Intervention & Protection	Innovative Research
Communication Information Systems & Networks Systems of systems Battlelab and Modelling & Simulation Cyber Research & Technology WG	Aerial Systems Ground Systems Naval Systems Ammunition Technology	Materials & Structures Technologies for Components and Modules Radio Frequency Sensors Technologies Electro-Optical Sensors Technologies CBRN Protection and Human Factors Guidance, Navigation & Control Energy WG

The detailed technical coverage of each group is posted on the **EDA WEBSITE**

EUROPEAN DEFENCE AGENCY

21

www.eda.europa.eu



What is a SRA?

- Strategic Research Agenda (SRA) is a document that provides an introduction to the technical field addressed by each CapTech/WG.
- It is linked from inception to the military capabilities and the technical base required to provide future solutions. The SRAs are intended to be used to provide strategic guidance for the R&T priorities addressed in the different CapTechs/WGs.

  In order to support the incorporation of new topics and technologies a three-step process is envisaged:
- technology watch
- assessment of technologies identified
- selection of the most promising technologies to be developed within the EDA framework.

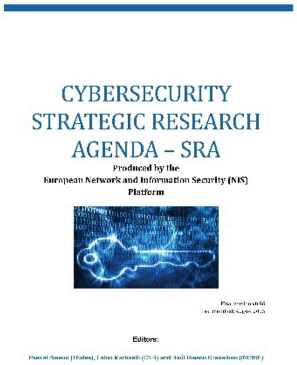
EUROPEAN DEFENCE AGENCY

22

www.eda.europa.eu

References

ECS
EUROPEAN CYBER SECURITY ORGANISATION
European Cybersecurity Strategic Research and Innovation Agenda (SRIA) for a contractual Public-Private Partnership (cPPP)


CYBERSECURITY STRATEGIC RESEARCH AGENDA – SRA
Produced by the
European Network and Information Security (NIS) Platform
Editors:

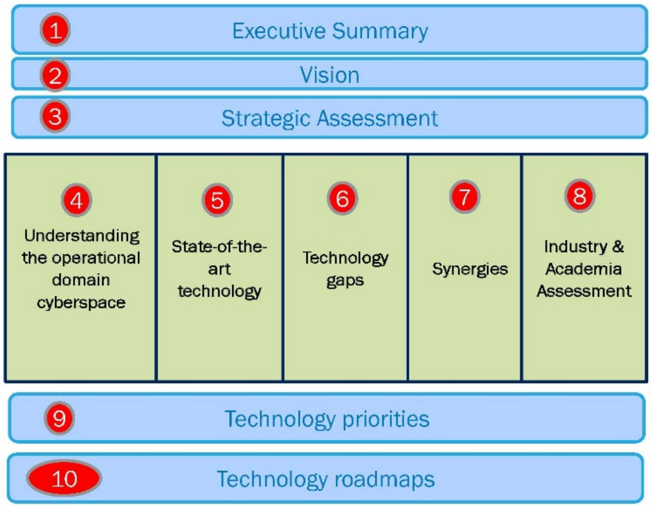
STRATEGIC RESEARCH & INNOVATION AGENDA
2017 UPDATE | VOLUME 1
ACARE
Advisory Council for Aviation Research and Innovation in Europe

http://www.ecs-org.eu/documents/ecs-cppp-sria.pdf

https://resilience.enisa.europa.eu/nis-platform/shared-documents/wg3-documents

http://www.acare4europe.org/sria

Cyber SRA content



1 Executive Summary

2 Vision

3 Strategic Assessment

4 Understanding the operational domain cyberspace

5 State-of-the-art technology

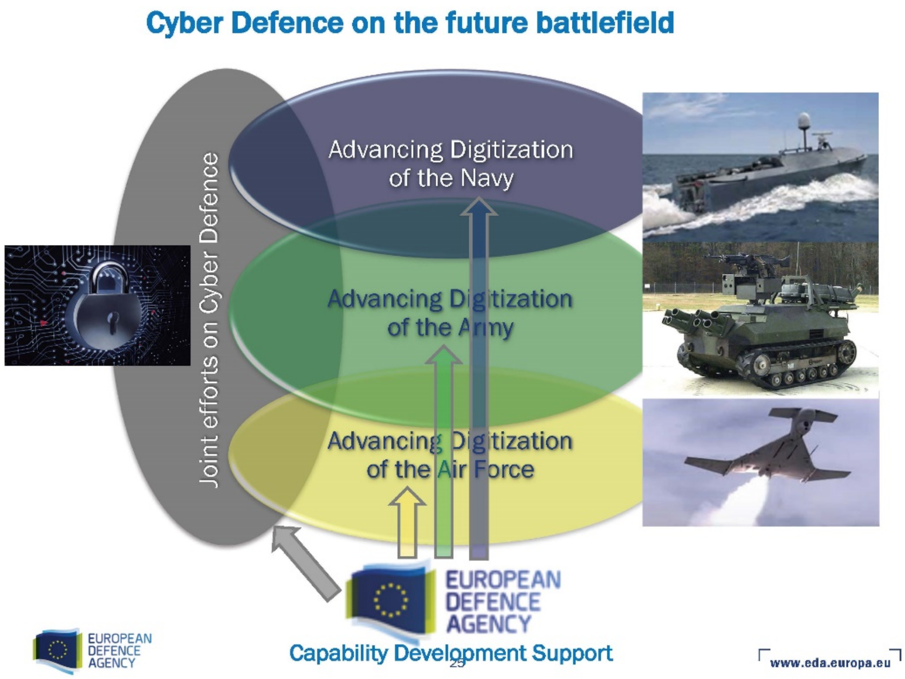
6 Technology gaps

7 Synergies

8 Industry & Academia Assessment

9 Technology priorities

10 Technology roadmaps



Cyber Defence on the future battlefield

Joint efforts on Cyber Defence

Advancing Digitization of the Navy

Advancing Digitization of the Army

Advancing Digitization of the Air Force

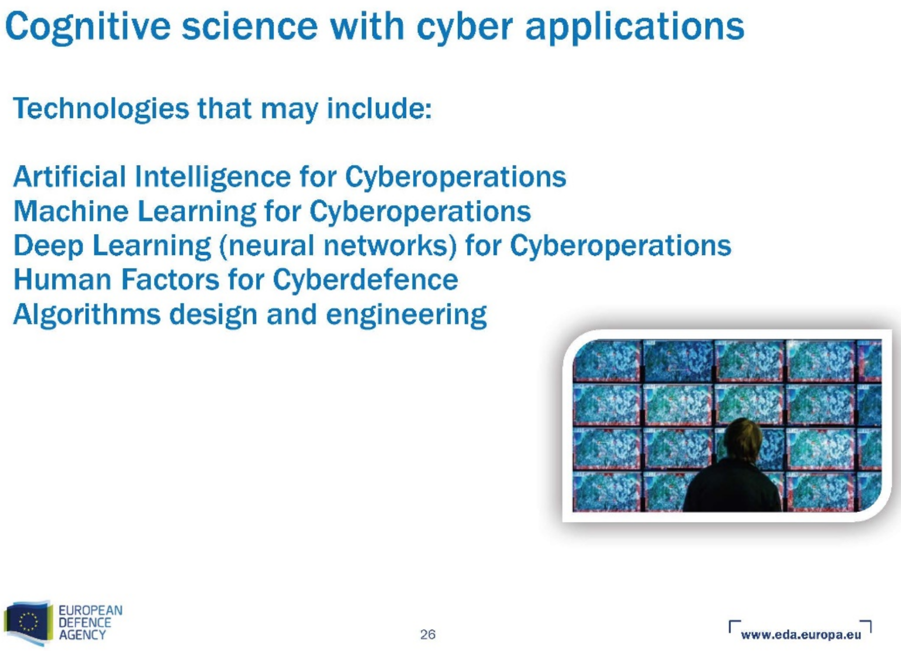
EUROPEAN DEFENCE AGENCY

Capability Development Support

EUROPEAN DEFENCE AGENCY

25

www.eda.europa.eu



Cognitive science with cyber applications

Technologies that may include:

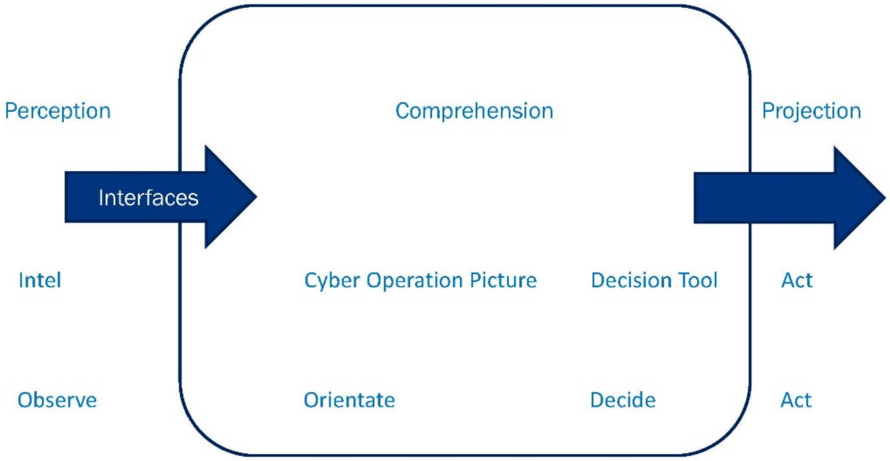
Artificial Intelligence for Cyberoperations
Machine Learning for Cyberoperations
Deep Learning (neural networks) for Cyberoperations
Human Factors for Cyberdefence
Algorithms design and engineering

EUROPEAN DEFENCE AGENCY

26

www.eda.europa.eu

Cyber Situation Awareness. Research Areas



Perception

Comprehension

Projection

Interfaces

Intel

Cyber Operation Picture

Decision Tool

Act

Observe

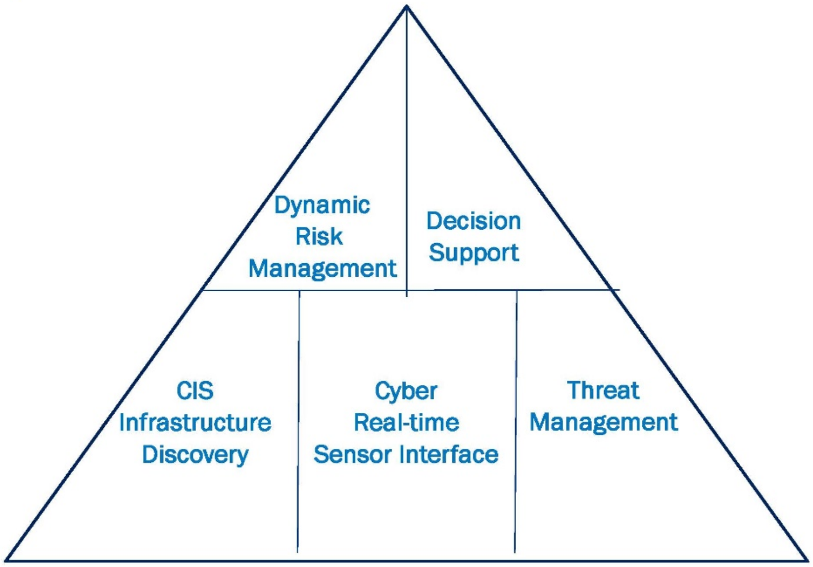
Orientate

Decide

Act

Source: M. R. Endsley, "Towards a Theory of Situation Awareness in Dynamic Systems", *Human Factors* 37 (1), 1995, p. 36.

Cyber Situation Awareness. Research Areas



Dynamic Risk Management

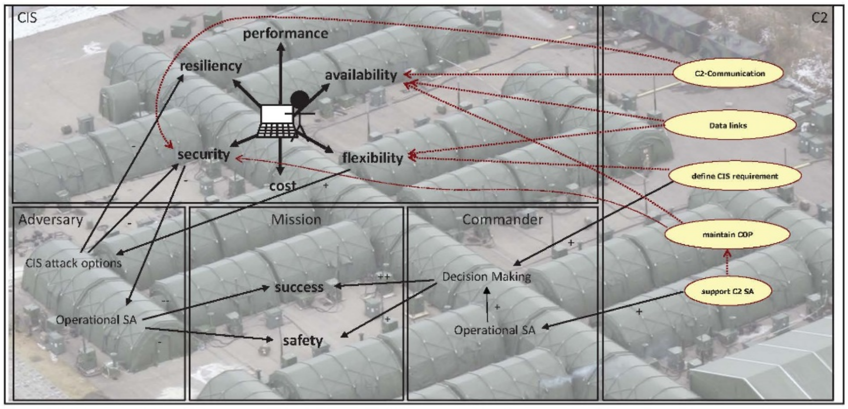
Decision Support

CIS Infrastructure Discovery

Cyber Real-time Sensor Interface

Threat Management

HF-HSI. CIS-use cases and mission goals on the operational level



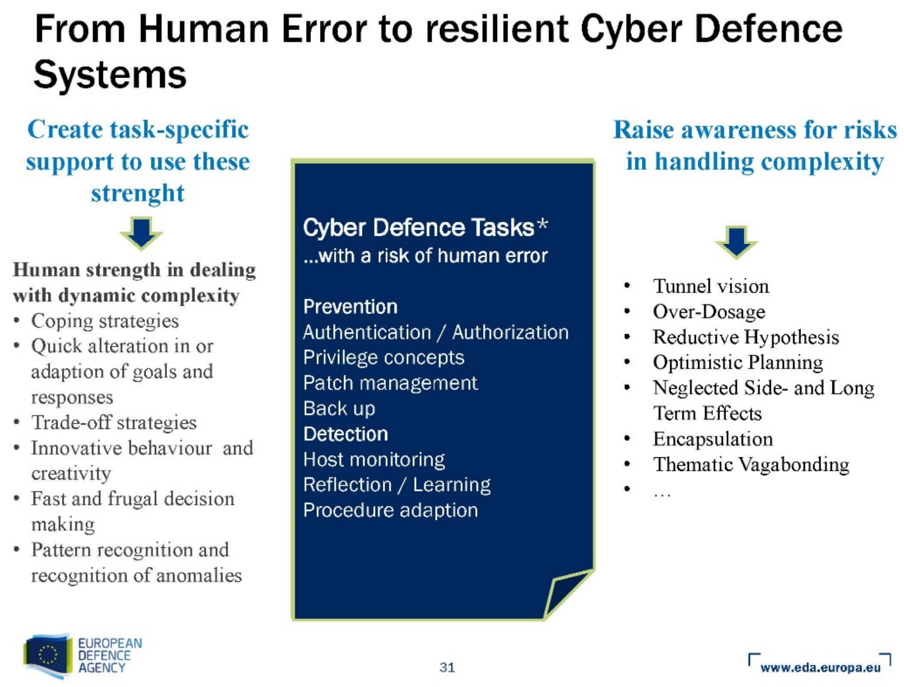
HF-Aspects Commander

Situational Awareness
How much of the Cyber situational picture should be part of the COP?

Decision under uncertainty
incident -> commander cannot be sure about security of information

Risk communication
need to communicate with the CD units, about risks in negotiating CIS requirements for the operational situation.

HF category	HF aspect for the commander in CD
Decision Making Human Error (K)	Mental model of cyber risks. What is the right abstraction for cyber risks in the mental model of a commander
	Decision under uncertainty. If an incident has been reported the commander cannot be sure about the integrity, confidentiality, and authenticity of information. This might indicate specific adversary actions.
Human error (SR)	Minor priority on the command level
Situational Awareness (SA)	Needs Situational Awareness about the common operational picture (COP) and the cyber situation. How much of the Cyber situational picture should be part of the COP?
	Attribution of cyber-attacks. Who should be attacked? What are the best counter-strikes to the attacks?
Attention	Attention on cyber incidents should completely be delegated to SOCs
Workload	Balancing workload on considering cyber threats in the decision making process. To what extend could the handling of cyber threats be delegated, if workload of the commander is high
Motivation	Minor priority on the command level
Communication	Need to communicate with the CD units, about risks in negotiating CIS requirements for the operational situation.
Trust	Trust in his staff to comply with security policies. Trust in technology: how secure is the system from a technological view. Trust influence how much uncertainty about system state the commander will assume in the decision making process.



From Human Error to resilient Cyber Defence Systems

Create task-specific support to use these strenght

Human strength in dealing with dynamic complexity

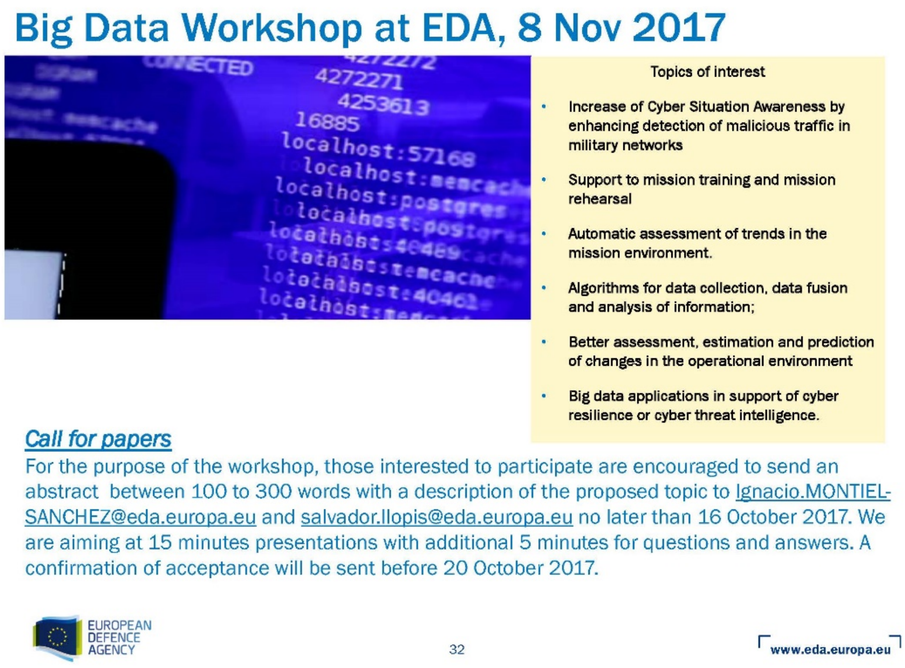
- Coping strategies
- Quick alteration in or adaption of goals and responses
- Trade-off strategies
- Innovative behaviour and creativity
- Fast and frugal decision making
- Pattern recognition and recognition of anomalies

Cyber Defence Tasks*
...with a risk of human error

Prevention
Authentication / Authorization
Privilege concepts
Patch management
Back up
Detection
Host monitoring
Reflection / Learning
Procedure adaption

Raise awareness for risks in handling complexity

- Tunnel vision
- Over-Dosage
- Reductive Hypothesis
- Optimistic Planning
- Neglected Side- and Long Term Effects
- Encapsulation
- Thematic Vagabonding
- …



Big Data Workshop at EDA, 8 Nov 2017

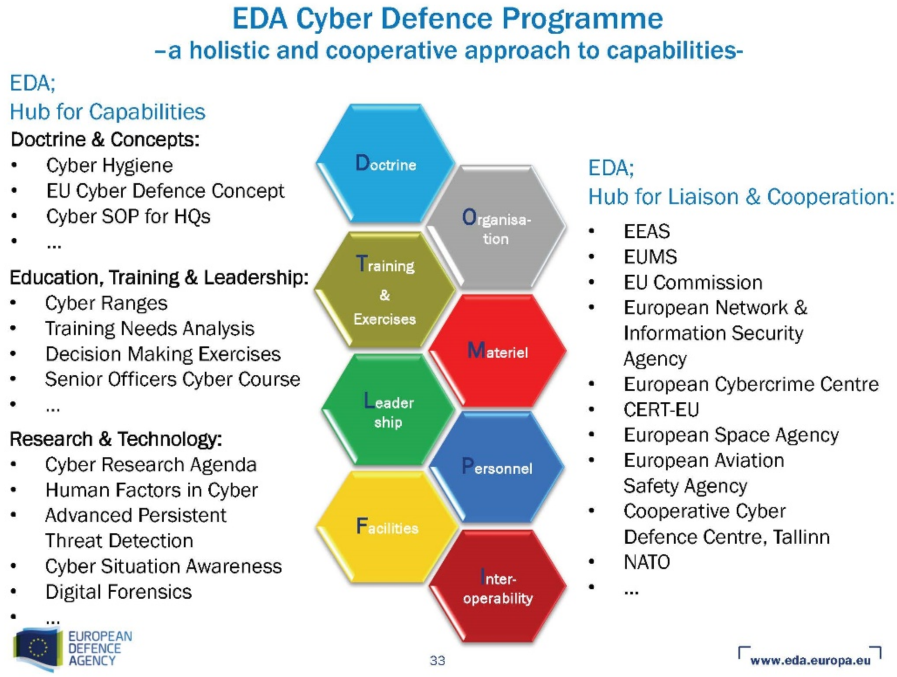
CONNECTED
4272271
4253613
16885
localhost:57168
localhost:memcache
localhost:postgres

Topics of interest

- Increase of Cyber Situation Awareness by enhancing detection of malicious traffic in military networks
- Support to mission training and mission rehearsal
- Automatic assessment of trends in the mission environment.
- Algorithms for data collection, data fusion and analysis of information;
- Better assessment, estimation and prediction of changes in the operational environment
- Big data applications in support of cyber resilience or cyber threat intelligence.

Call for papers

For the purpose of the workshop, those interested to participate are encouraged to send an abstract between 100 to 300 words with a description of the proposed topic to Ignacio.MONTIEL-SANCHEZ@eda.europa.eu and salvador.llopis@eda.europa.eu no later than 16 October 2017. We are aiming at 15 minutes presentations with additional 5 minutes for questions and answers. A confirmation of acceptance will be sent before 20 October 2017.



EDA Cyber Defence Programme
–a holistic and cooperative approach to capabilities-

EDA;
Hub for Capabilities

Doctrine & Concepts:
- Cyber Hygiene
- EU Cyber Defence Concept
- Cyber SOP for HQs
- …

Education, Training & Leadership:
- Cyber Ranges
- Training Needs Analysis
- Decision Making Exercises
- Senior Officers Cyber Course
- …

Research & Technology:
- Cyber Research Agenda
- Human Factors in Cyber
- Advanced Persistent Threat Detection
- Cyber Situation Awareness
- Digital Forensics
- …

Doctrine

Organisa-tion

Training & Exercises

Materiel

Leader ship

Personnel

Facilities

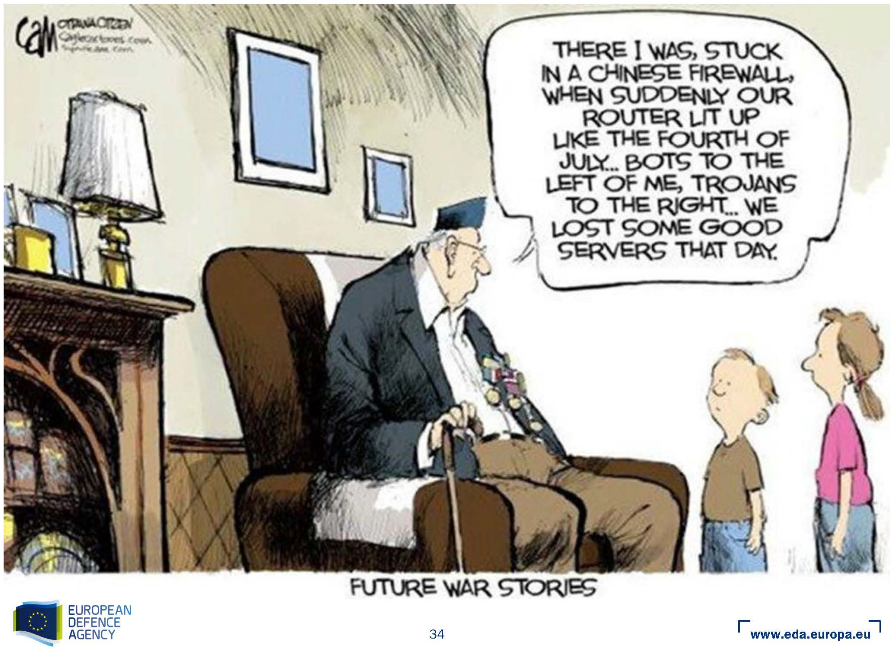
Inter-operability

EDA;
Hub for Liaison & Cooperation:
- EEAS
- EUMS
- EU Commission
- European Network & Information Security Agency
- European Cybercrime Centre
- CERT-EU
- European Space Agency
- European Aviation Safety Agency
- Cooperative Cyber Defence Centre, Tallinn
- NATO
- …

EUROPEAN DEFENCE AGENCY

33

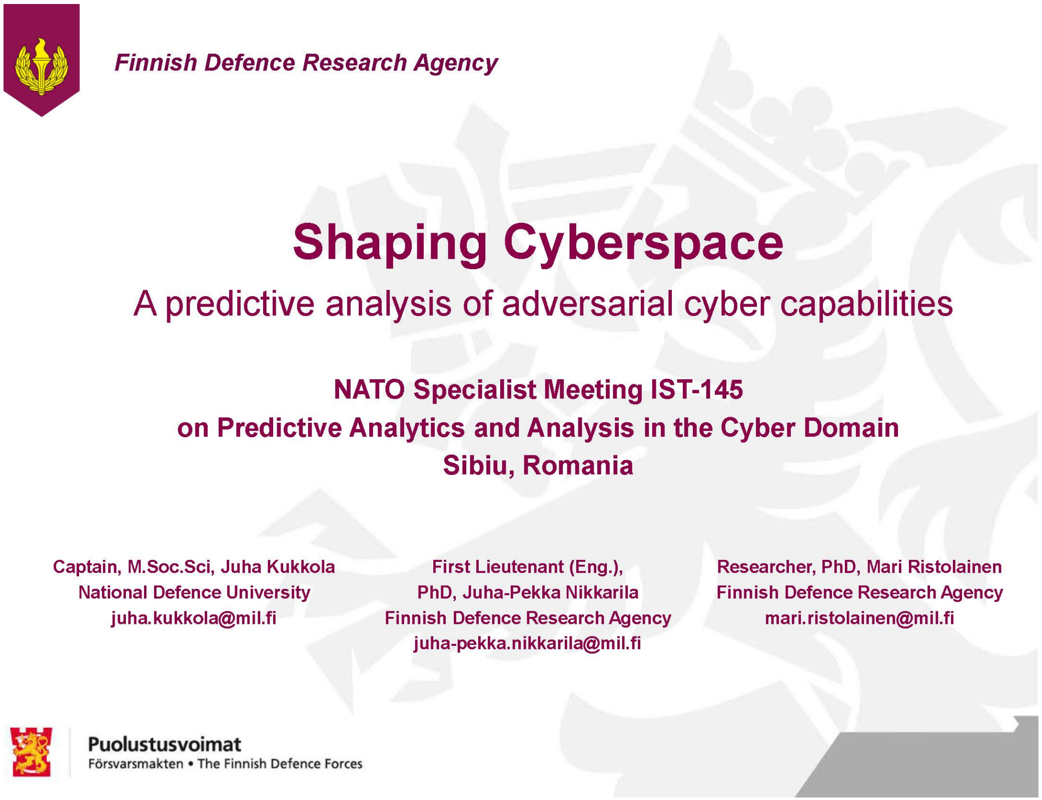
www.eda.europa.eu


CAM OTTAWA CITIZEN
THERE I WAS, STUCK IN A CHINESE FIREWALL, WHEN SUDDENLY OUR ROUTER LIT UP LIKE THE FOURTH OF JULY... BOTS TO THE LEFT OF ME, TROJANS TO THE RIGHT... WE LOST SOME GOOD SERVERS THAT DAY.
FUTURE WAR STORIES

EUROPEAN DEFENCE AGENCY

34

www.eda.europa.eu



Finnish Defence Research Agency

Shaping Cyberspace

A predictive analysis of adversarial cyber capabilities

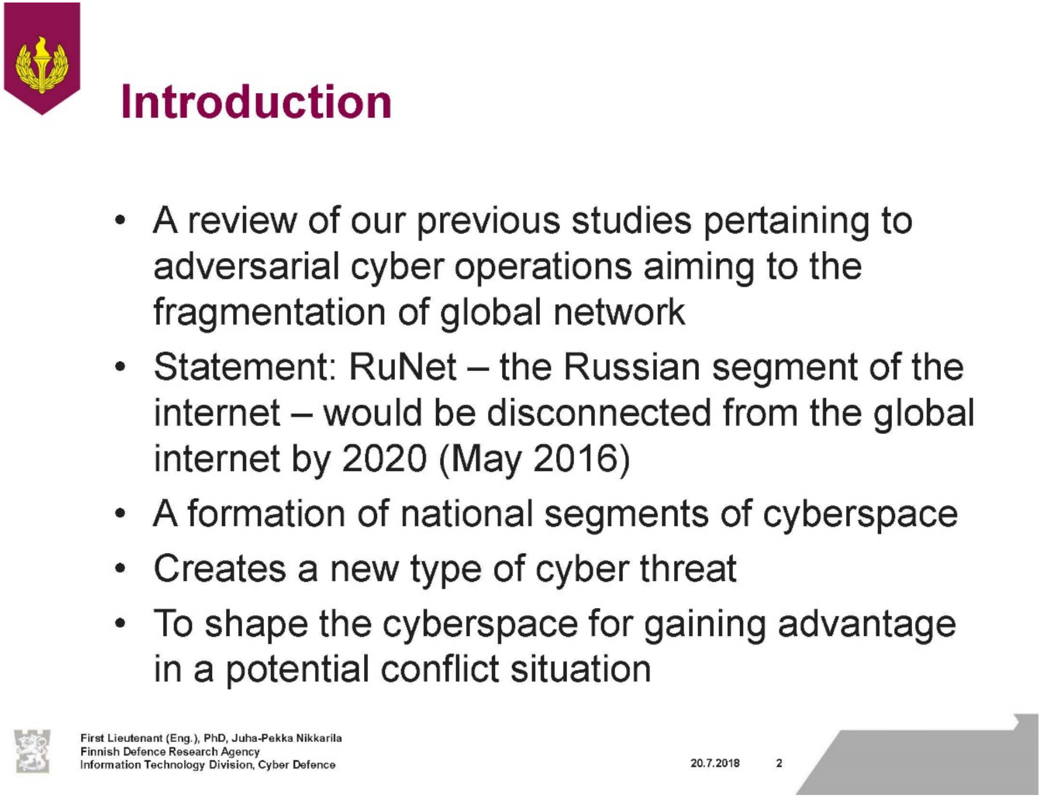
NATO Specialist Meeting IST-145
on Predictive Analytics and Analysis in the Cyber Domain
Sibiu, Romania

Captain, M.Soc.Sci, Juha Kukkola
National Defence University
juha.kukkola@mil.fi

First Lieutenant (Eng.),
PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
juha-pekka.nikkarila@mil.fi

Researcher, PhD, Mari Ristolainen
Finnish Defence Research Agency
mari.ristolainen@mil.fi

Puolustusvoimat
Försvarsmakten • The Finnish Defence Forces

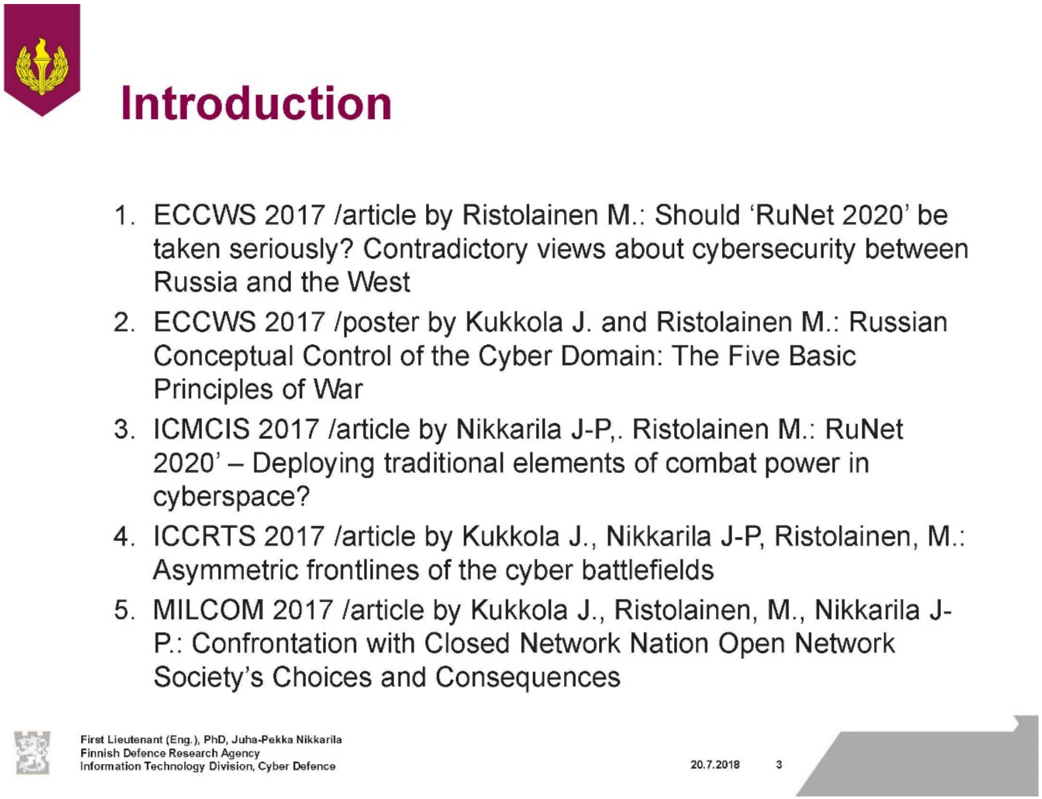


Introduction

- A review of our previous studies pertaining to adversarial cyber operations aiming to the fragmentation of global network
- Statement: RuNet – the Russian segment of the internet – would be disconnected from the global internet by 2020 (May 2016)
- A formation of national segments of cyberspace
- Creates a new type of cyber threat
- To shape the cyberspace for gaining advantage in a potential conflict situation

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 2

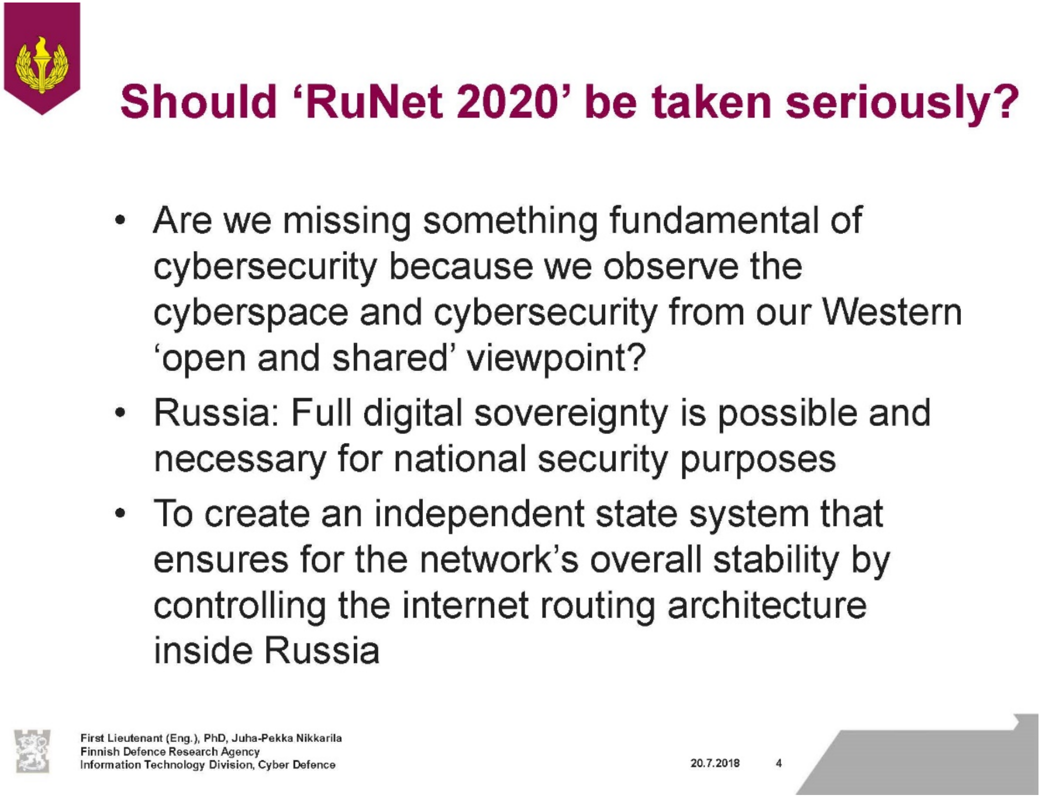


Introduction

1. ECCWS 2017 /article by Ristolainen M.: Should 'RuNet 2020' be taken seriously? Contradictory views about cybersecurity between Russia and the West
2. ECCWS 2017 /poster by Kukkola J. and Ristolainen M.: Russian Conceptual Control of the Cyber Domain: The Five Basic Principles of War
3. ICMCIS 2017 /article by Nikkarila J-P,. Ristolainen M.: RuNet 2020' – Deploying traditional elements of combat power in cyberspace?
4. ICCRTS 2017 /article by Kukkola J., Nikkarila J-P, Ristolainen, M.: Asymmetric frontlines of the cyber battlefields
5. MILCOM 2017 /article by Kukkola J., Ristolainen, M., Nikkarila J-P.: Confrontation with Closed Network Nation Open Network Society's Choices and Consequences

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 3

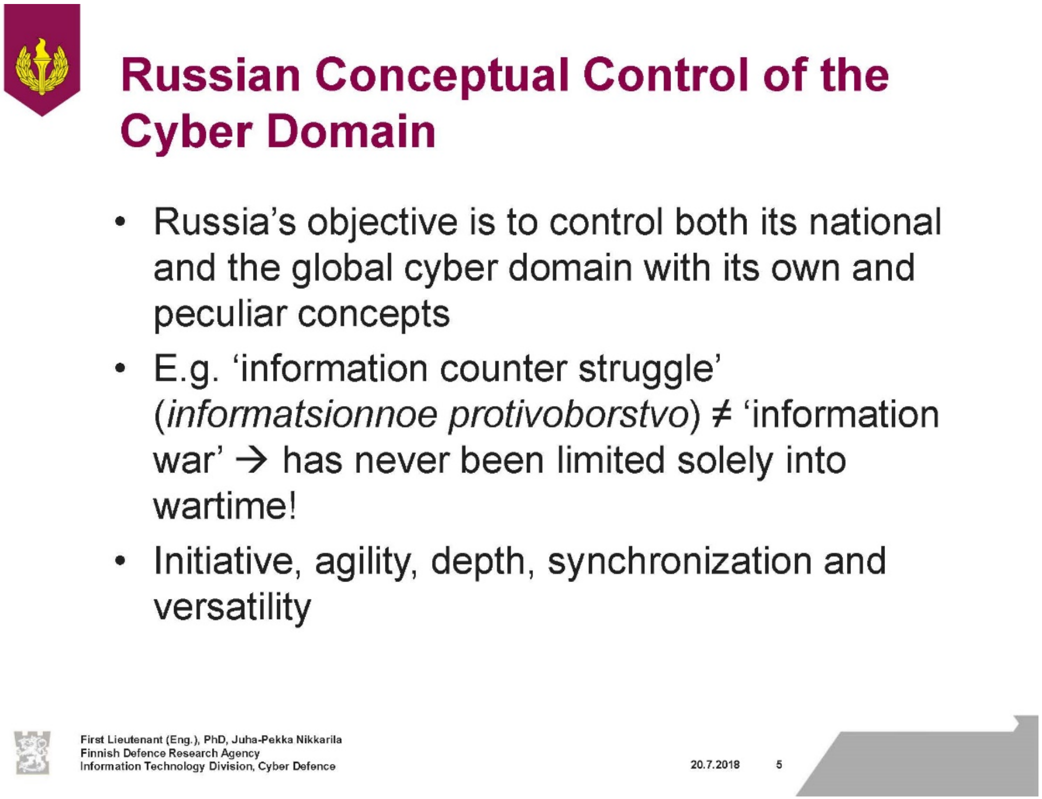


Should 'RuNet 2020' be taken seriously?

- Are we missing something fundamental of cybersecurity because we observe the cyberspace and cybersecurity from our Western 'open and shared' viewpoint?
- Russia: Full digital sovereignty is possible and necessary for national security purposes
- To create an independent state system that ensures for the network's overall stability by controlling the internet routing architecture inside Russia

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 4

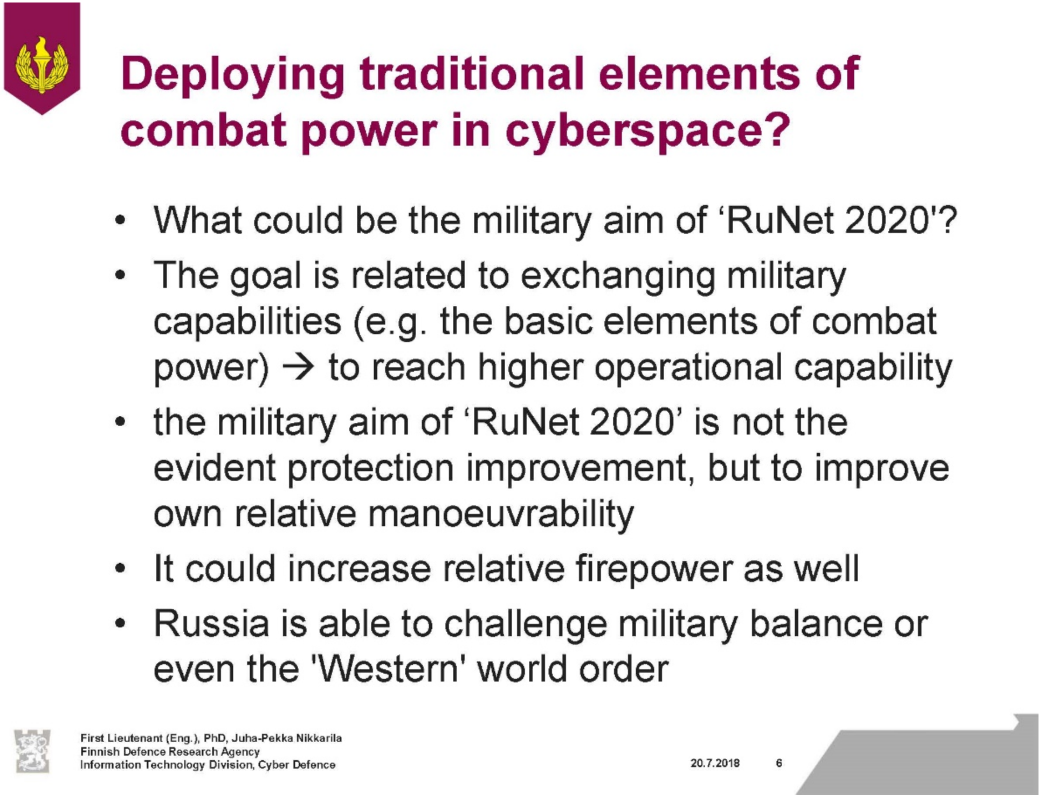


Russian Conceptual Control of the Cyber Domain

- Russia's objective is to control both its national and the global cyber domain with its own and peculiar concepts
- E.g. 'information counter struggle' (*informatsionnoe protivoborstvo*) ≠ 'information war' → has never been limited solely into wartime!
- Initiative, agility, depth, synchronization and versatility

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 5

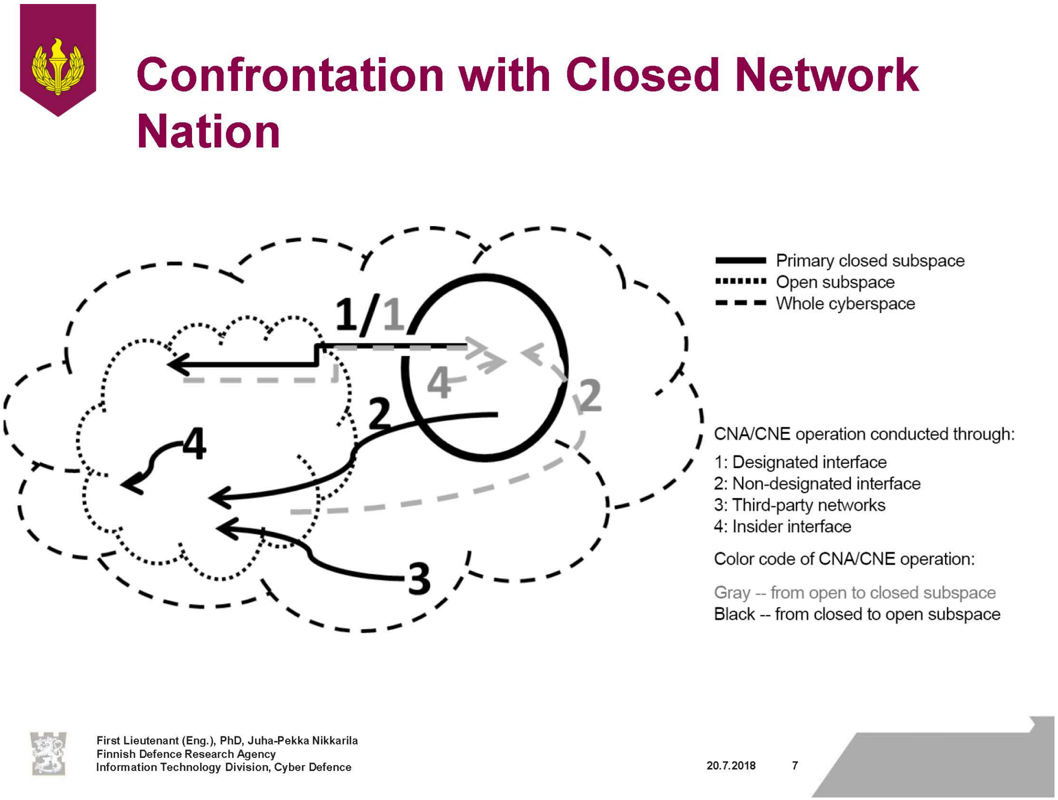


Deploying traditional elements of combat power in cyberspace?

- What could be the military aim of 'RuNet 2020'?
- The goal is related to exchanging military capabilities (e.g. the basic elements of combat power) → to reach higher operational capability
- the military aim of 'RuNet 2020' is not the evident protection improvement, but to improve own relative manoeuvrability
- It could increase relative firepower as well
- Russia is able to challenge military balance or even the 'Western' world order

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 6



Confrontation with Closed Network Nation

▬ Primary closed subspace
┈ Open subspace
╌ Whole cyberspace

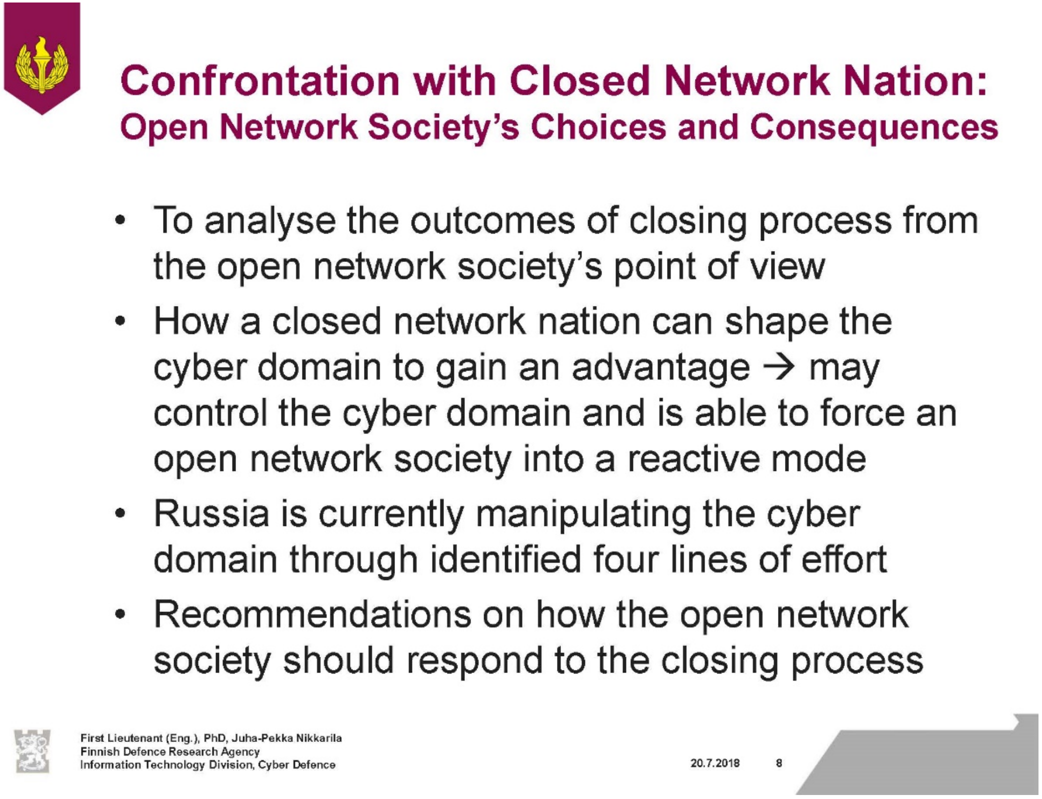
CNA/CNE operation conducted through:
1: Designated interface
2: Non-designated interface
3: Third-party networks
4: Insider interface

Color code of CNA/CNE operation:

Gray -- from open to closed subspace
Black -- from closed to open subspace

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 7

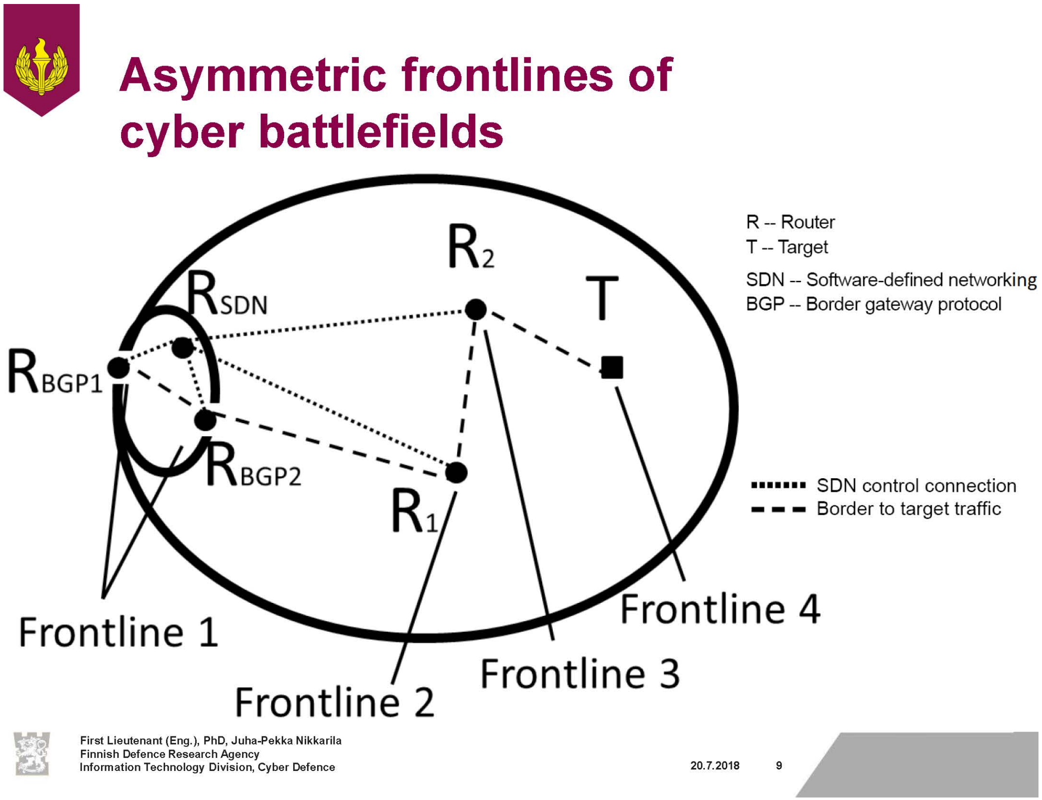


Confrontation with Closed Network Nation:
Open Network Society's Choices and Consequences

- To analyse the outcomes of closing process from the open network society's point of view
- How a closed network nation can shape the cyber domain to gain an advantage → may control the cyber domain and is able to force an open network society into a reactive mode
- Russia is currently manipulating the cyber domain through identified four lines of effort
- Recommendations on how the open network society should respond to the closing process

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 8



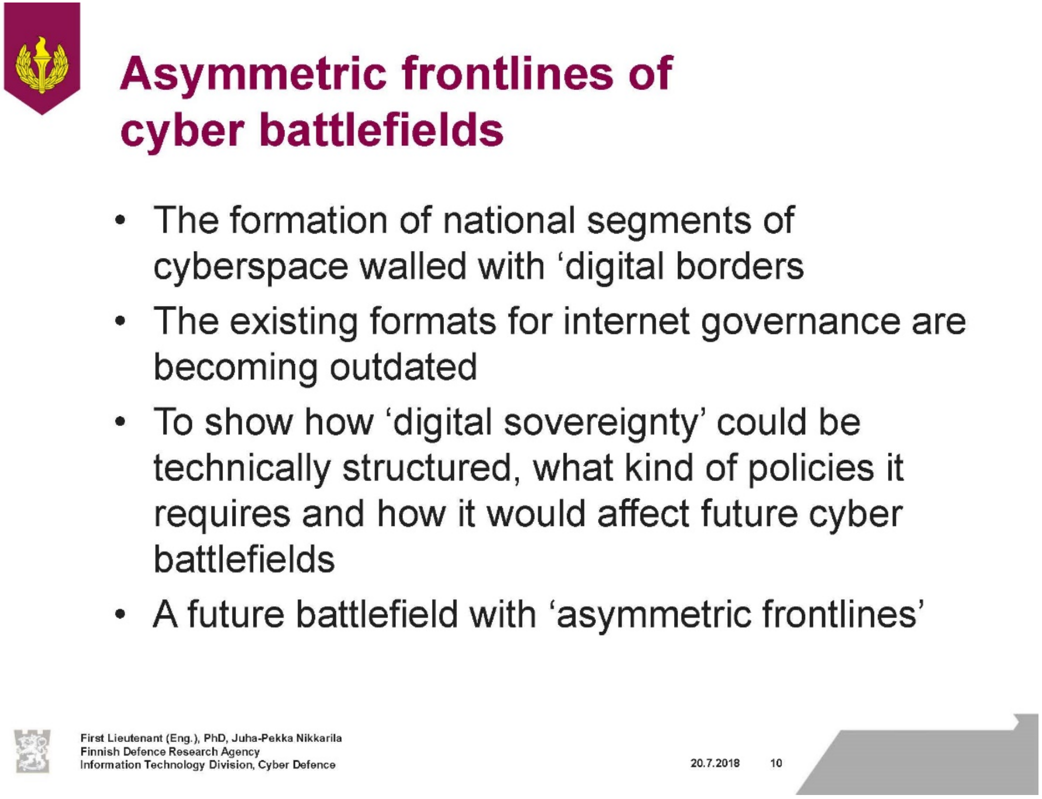
Asymmetric frontlines of cyber battlefields

R -- Router
T -- Target
SDN -- Software-defined networking
BGP -- Border gateway protocol

▪▪▪▪▪▪ SDN control connection
– – – Border to target traffic

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 9

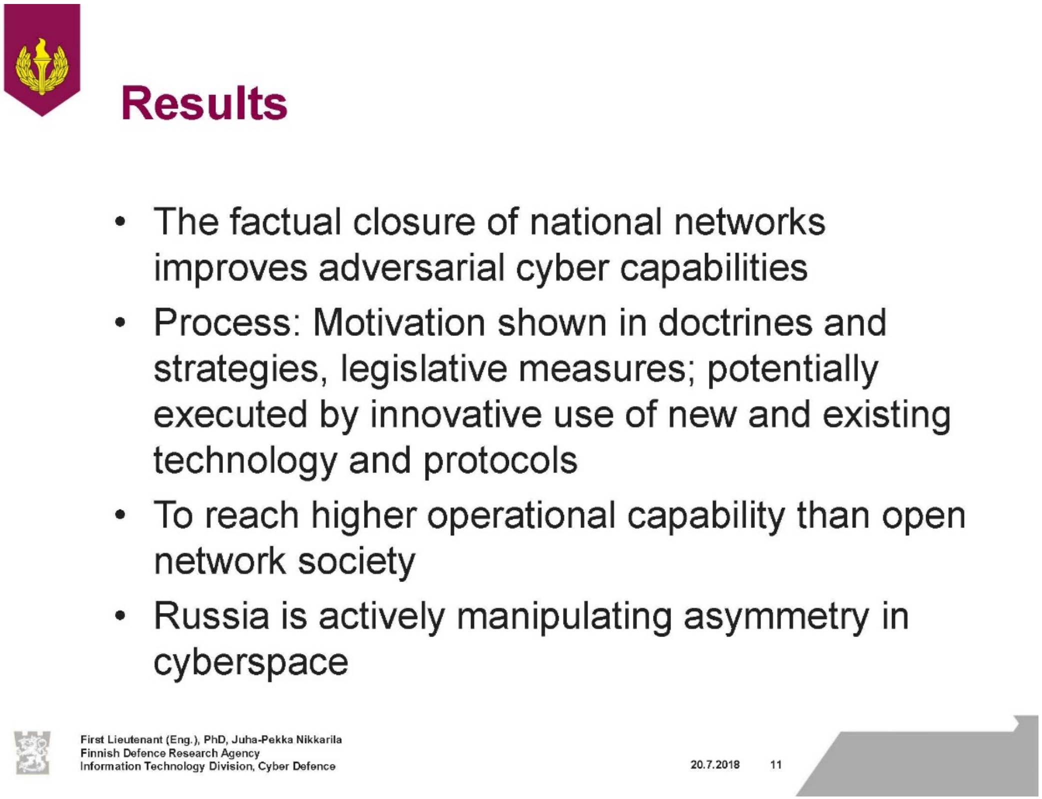


Asymmetric frontlines of cyber battlefields

- The formation of national segments of cyberspace walled with ‘digital borders
- The existing formats for internet governance are becoming outdated
- To show how ‘digital sovereignty’ could be technically structured, what kind of policies it requires and how it would affect future cyber battlefields
- A future battlefield with ‘asymmetric frontlines’

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 10

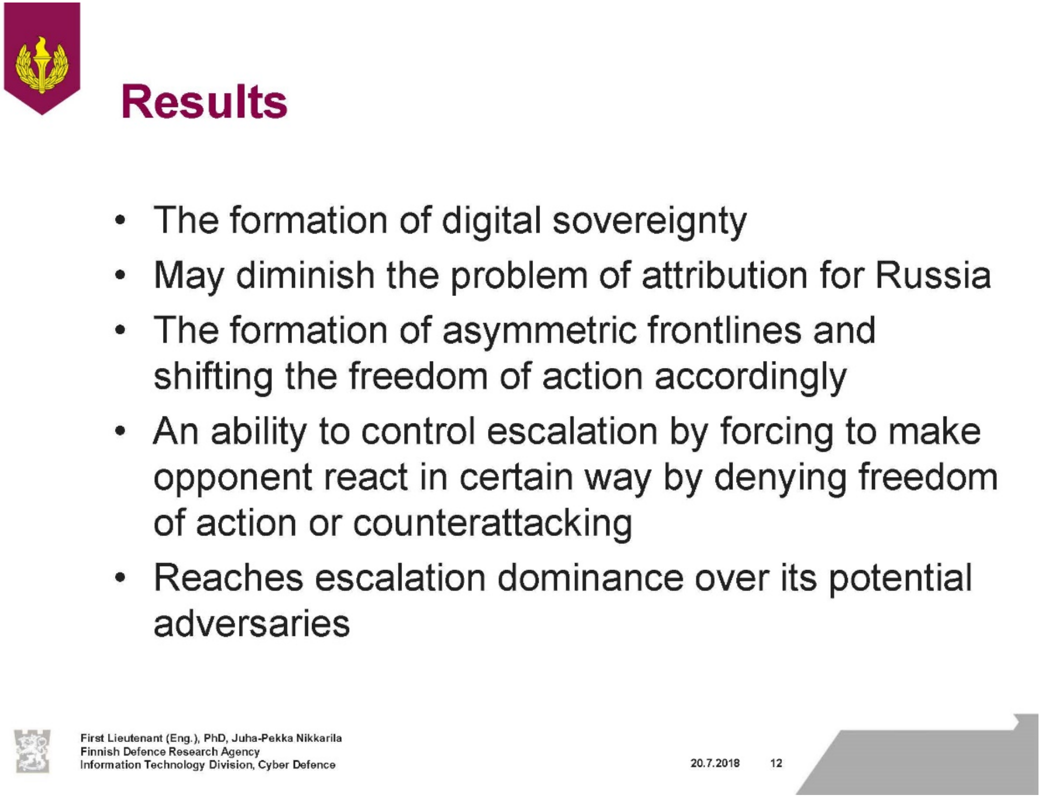


Results

- The factual closure of national networks improves adversarial cyber capabilities
- Process: Motivation shown in doctrines and strategies, legislative measures; potentially executed by innovative use of new and existing technology and protocols
- To reach higher operational capability than open network society
- Russia is actively manipulating asymmetry in cyberspace

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 11

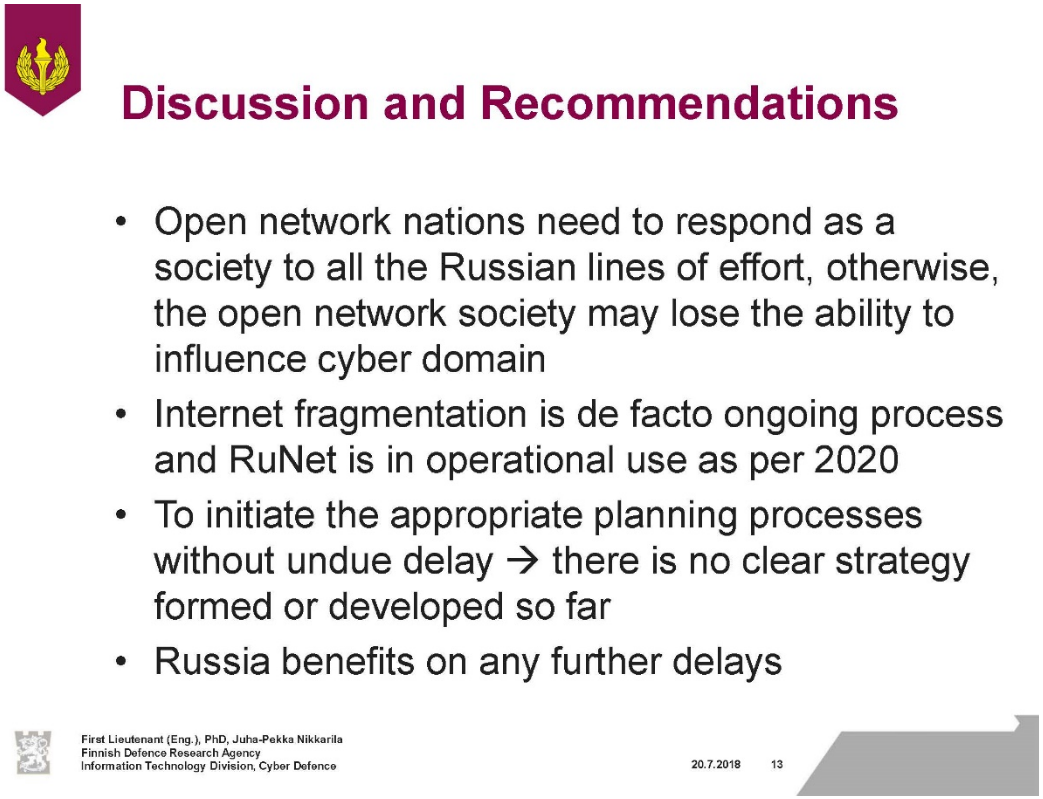


Results

- The formation of digital sovereignty
- May diminish the problem of attribution for Russia
- The formation of asymmetric frontlines and shifting the freedom of action accordingly
- An ability to control escalation by forcing to make opponent react in certain way by denying freedom of action or counterattacking
- Reaches escalation dominance over its potential adversaries

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 12

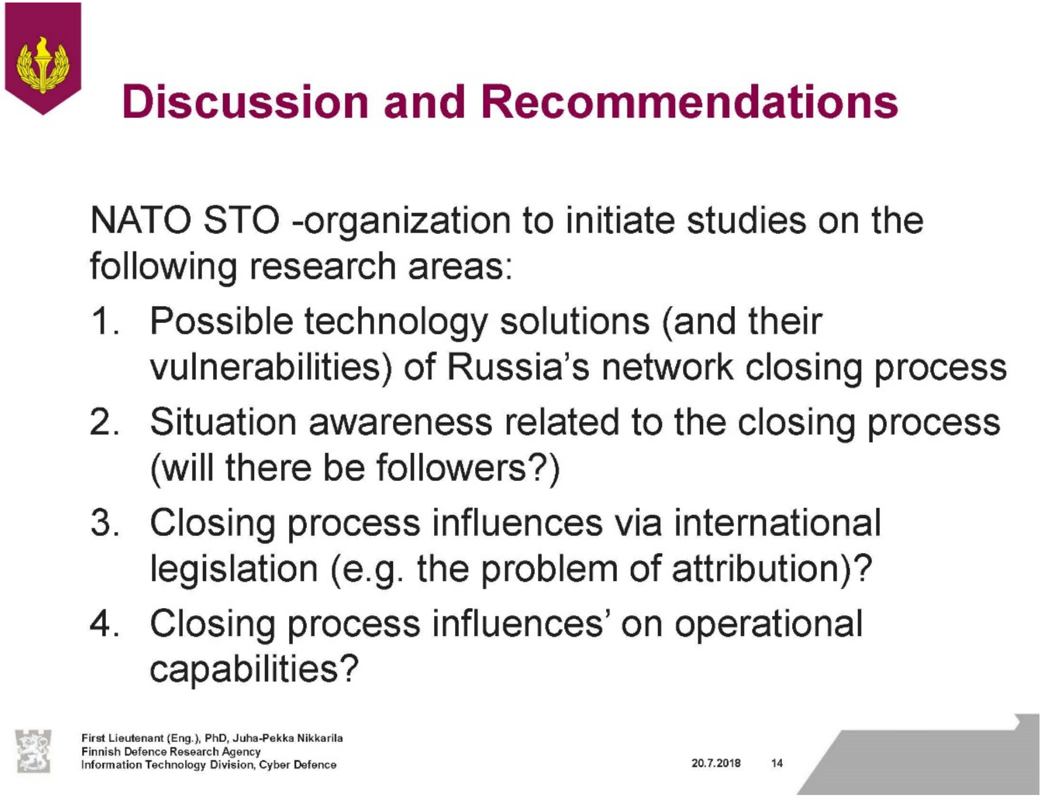


Discussion and Recommendations

- Open network nations need to respond as a society to all the Russian lines of effort, otherwise, the open network society may lose the ability to influence cyber domain
- Internet fragmentation is de facto ongoing process and RuNet is in operational use as per 2020
- To initiate the appropriate planning processes without undue delay → there is no clear strategy formed or developed so far
- Russia benefits on any further delays

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 13



Discussion and Recommendations

NATO STO -organization to initiate studies on the following research areas:

1. Possible technology solutions (and their vulnerabilities) of Russia's network closing process
2. Situation awareness related to the closing process (will there be followers?)
3. Closing process influences via international legislation (e.g. the problem of attribution)?
4. Closing process influences' on operational capabilities?

First Lieutenant (Eng.), PhD, Juha-Pekka Nikkarila
Finnish Defence Research Agency
Information Technology Division, Cyber Defence

20.7.2018 14

UNCLASSIFIED

U.S. ARMY
U.S. ARMY RDECOM
ARL

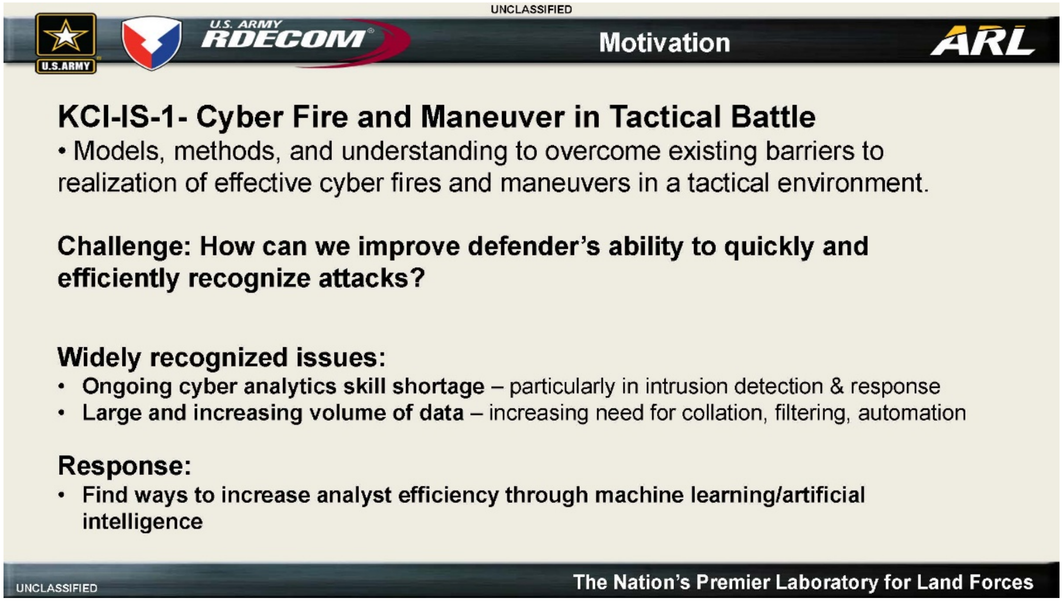
Intrusion Detection and Prevention System Alert Prioritization through Supervised Learning

Mr. Gregory Shearer (ARL - CTR), **Dr. Nandi Leslie** (ARL - CTR), **Mr. Paul Ritchey** (ARL - CTR), **Dr. Frederica Nelson** (ARL), **Dr. Tracy Braun** (ARL)

Network Security Branch, ARL

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY
U.S. ARMY RDECOM
ARL

Motivation

KCI-IS-1- Cyber Fire and Maneuver in Tactical Battle

• Models, methods, and understanding to overcome existing barriers to realization of effective cyber fires and maneuvers in a tactical environment.

Challenge: How can we improve defender's ability to quickly and efficiently recognize attacks?

Widely recognized issues:

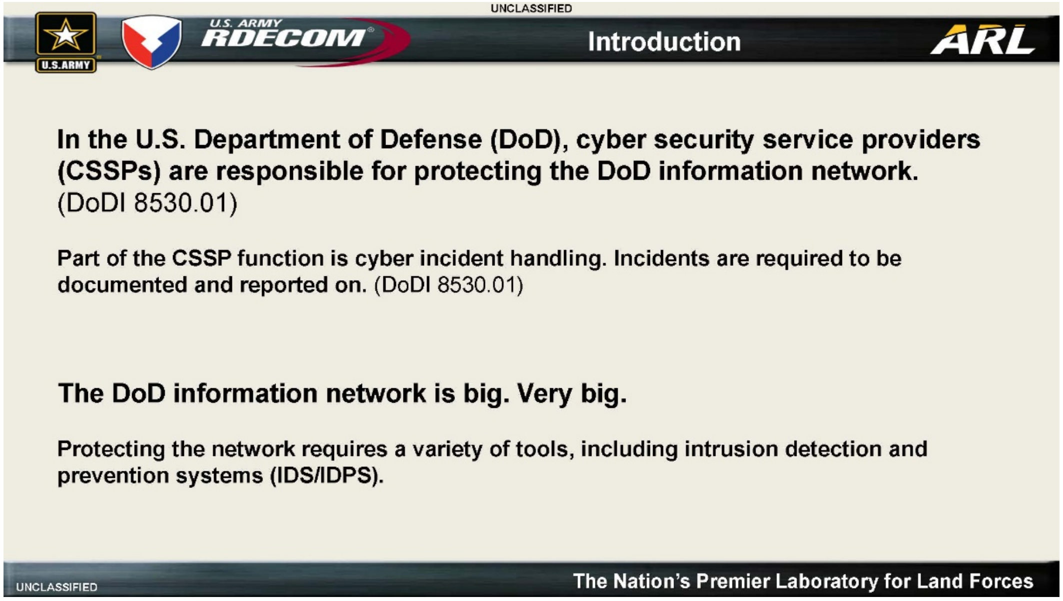
- **Ongoing cyber analytics skill shortage** – particularly in intrusion detection & response
- **Large and increasing volume of data** – increasing need for collation, filtering, automation

Response:

- **Find ways to increase analyst efficiency through machine learning/artificial intelligence**

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Introduction

ARL

In the U.S. Department of Defense (DoD), cyber security service providers (CSSPs) are responsible for protecting the DoD information network. (DoDI 8530.01)

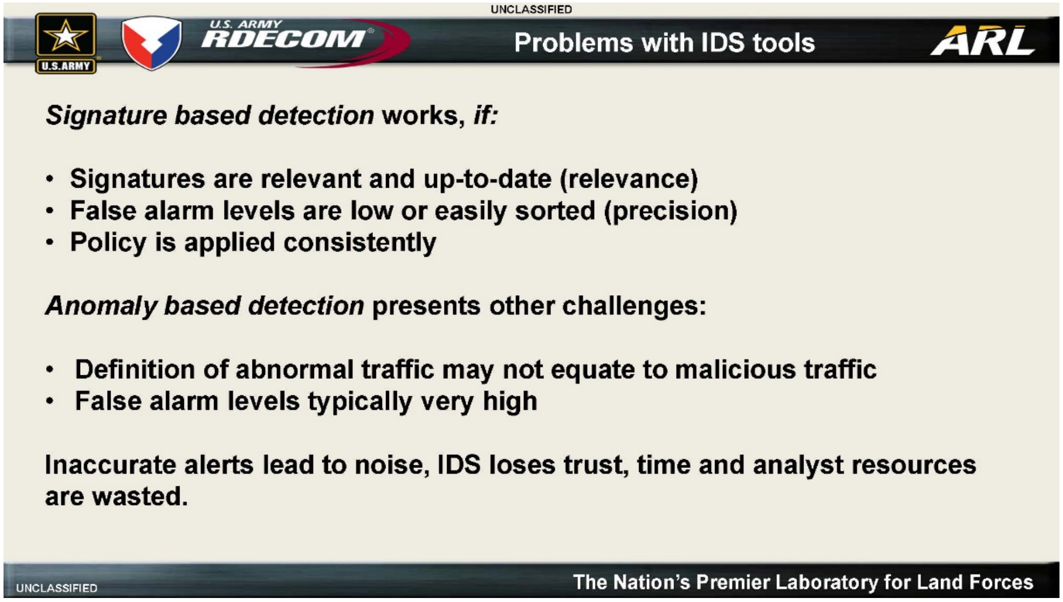
Part of the CSSP function is cyber incident handling. Incidents are required to be documented and reported on. (DoDI 8530.01)

The DoD information network is big. Very big.

Protecting the network requires a variety of tools, including intrusion detection and prevention systems (IDS/IDPS).

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Problems with IDS tools

ARL

Signature based detection* works, *if:

- **Signatures are relevant and up-to-date (relevance)**
- **False alarm levels are low or easily sorted (precision)**
- **Policy is applied consistently**

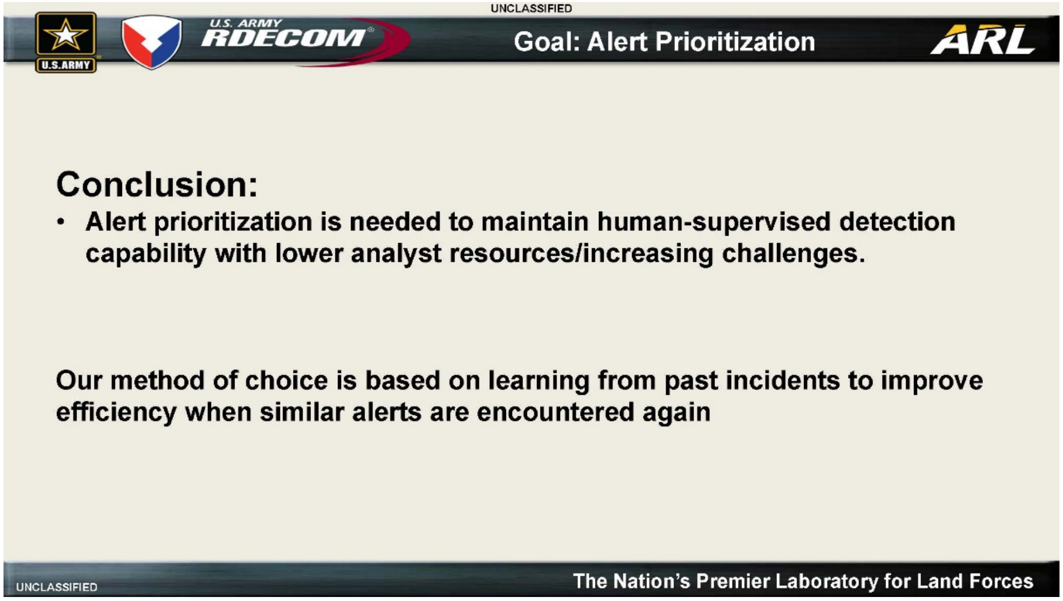
***Anomaly based detection* presents other challenges:**

- **Definition of abnormal traffic may not equate to malicious traffic**
- **False alarm levels typically very high**

Inaccurate alerts lead to noise, IDS loses trust, time and analyst resources are wasted.

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Goal: Alert Prioritization

ARL

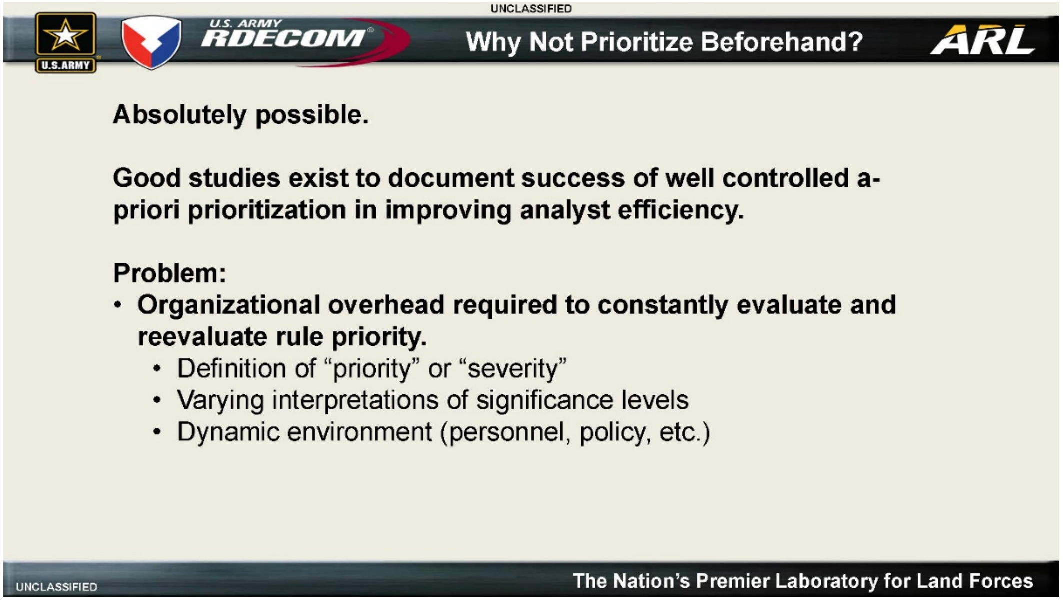
Conclusion:

- Alert prioritization is needed to maintain human-supervised detection capability with lower analyst resources/increasing challenges.

Our method of choice is based on learning from past incidents to improve efficiency when similar alerts are encountered again

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Why Not Prioritize Beforehand?

ARL

Absolutely possible.

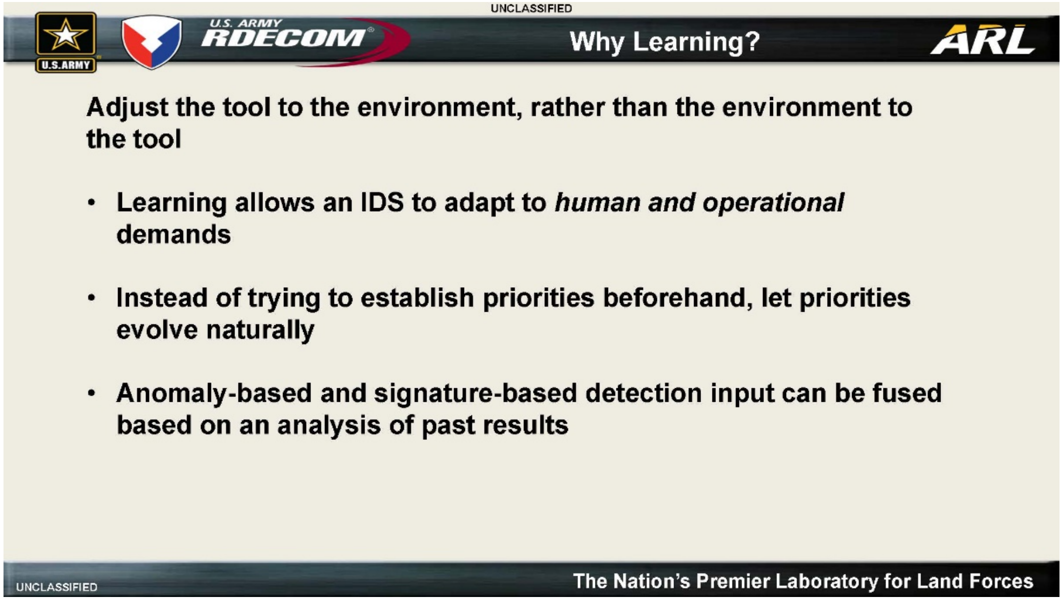
Good studies exist to document success of well controlled a-priori prioritization in improving analyst efficiency.

Problem:

- **Organizational overhead required to constantly evaluate and reevaluate rule priority.**
  - Definition of “priority” or “severity”
  - Varying interpretations of significance levels
  - Dynamic environment (personnel, policy, etc.)

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Why Learning?

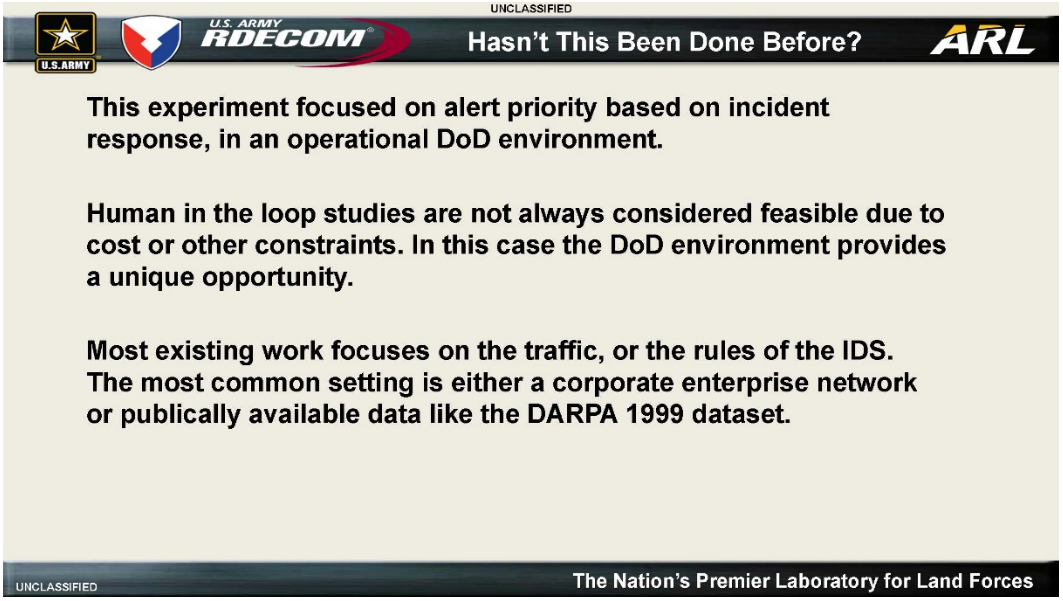
ARL

Adjust the tool to the environment, rather than the environment to the tool

- Learning allows an IDS to adapt to *human and operational* demands
- Instead of trying to establish priorities beforehand, let priorities evolve naturally
- Anomaly-based and signature-based detection input can be fused based on an analysis of past results

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

Hasn't This Been Done Before?

ARL

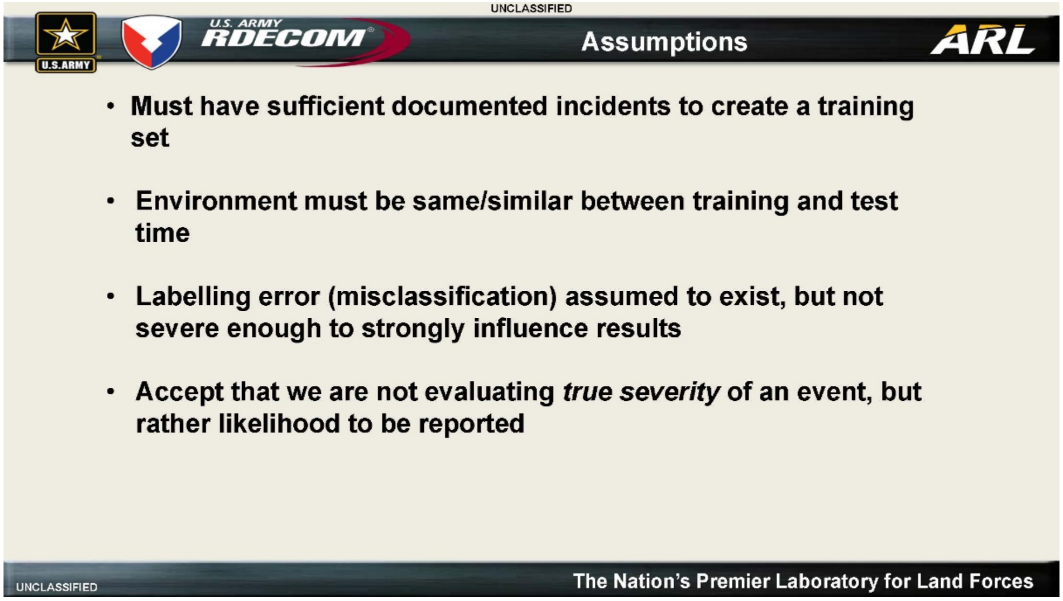
This experiment focused on alert priority based on incident response, in an operational DoD environment.

Human in the loop studies are not always considered feasible due to cost or other constraints. In this case the DoD environment provides a unique opportunity.

Most existing work focuses on the traffic, or the rules of the IDS. The most common setting is either a corporate enterprise network or publically available data like the DARPA 1999 dataset.

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



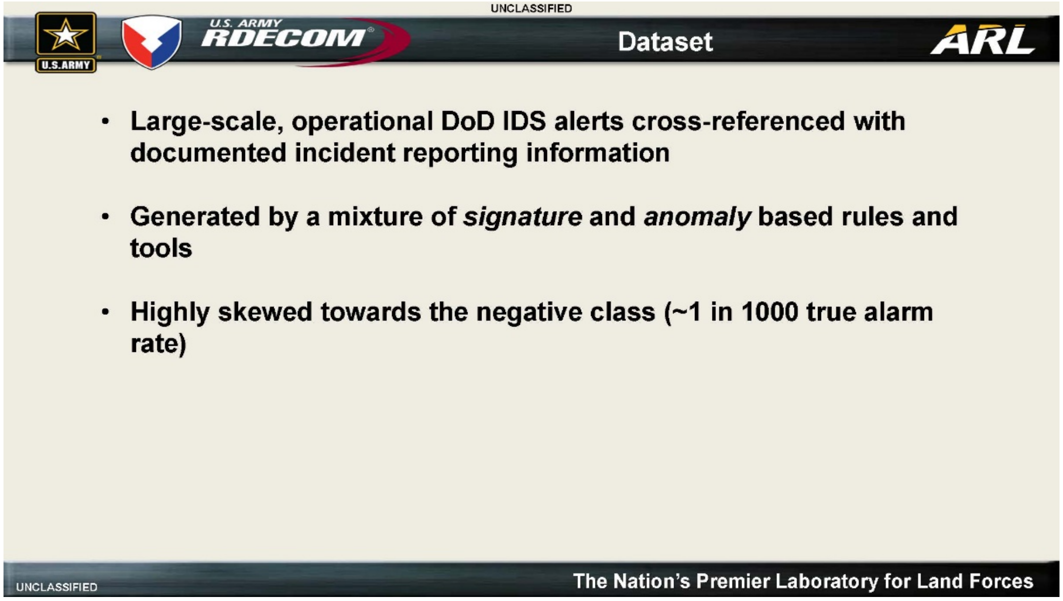
UNCLASSIFIED

Assumptions

- Must have sufficient documented incidents to create a training set
- Environment must be same/similar between training and test time
- Labelling error (misclassification) assumed to exist, but not severe enough to strongly influence results
- Accept that we are not evaluating ***true severity*** of an event, but rather likelihood to be reported

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

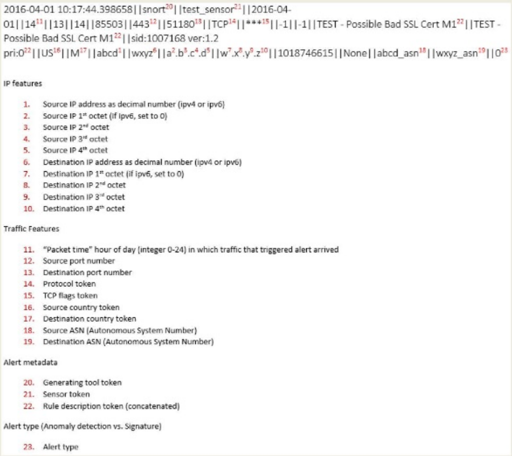
Dataset

- Large-scale, operational DoD IDS alerts cross-referenced with documented incident reporting information
- Generated by a mixture of ***signature*** and ***anomaly*** based rules and tools
- Highly skewed towards the negative class (~1 in 1000 true alarm rate)

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces

Data Transformation



2016-04-01 10:17:44.398658||snort^{20}||test_sensor^{21}||2016-04-01||14^{11}||13||14||85503||443^{12}||51180^{13}||TCP^{14}||***^{15}||-1||-1||TEST - Possible Bad SSL Cert M1^{22}||TEST - Possible Bad SSL Cert M1^{22}||sid:1007168 ver:1.2 pri:0^{22}||US^{16}||M^{17}||abcd^{1}||wxyz^{6}||a^{2}.b^{3}.c^{4}.d^{5}||w^{7}.x^{8}.y^{9}.z^{10}||1018746615||None||abcd_asn^{18}||wxyz_asn^{19}||0^{23}

IP features

1. Source IP address as decimal number (ipv4 or ipv6)
2. Source IP 1st octet (if ipv6, set to 0)
3. Source IP 2nd octet
4. Source IP 3rd octet
5. Source IP 4th octet
6. Destination IP address as decimal number (ipv4 or ipv6)
7. Destination IP 1st octet (if ipv6, set to 0)
8. Destination IP 2nd octet
9. Destination IP 3rd octet
10. Destination IP 4th octet

Traffic Features

11. "Packet time" hour of day (integer 0-24) in which traffic that triggered alert arrived
12. Source port number
13. Destination port number
14. Protocol token
15. TCP flags token
16. Source country token
17. Destination country token
18. Source ASN (Autonomous System Number)
19. Destination ASN (Autonomous System Number)

Alert metadata

20. Generating tool token
21. Sensor token
22. Rule description token (concatenated)

Alert type (Anomaly detection vs. Signature)

23. Alert type

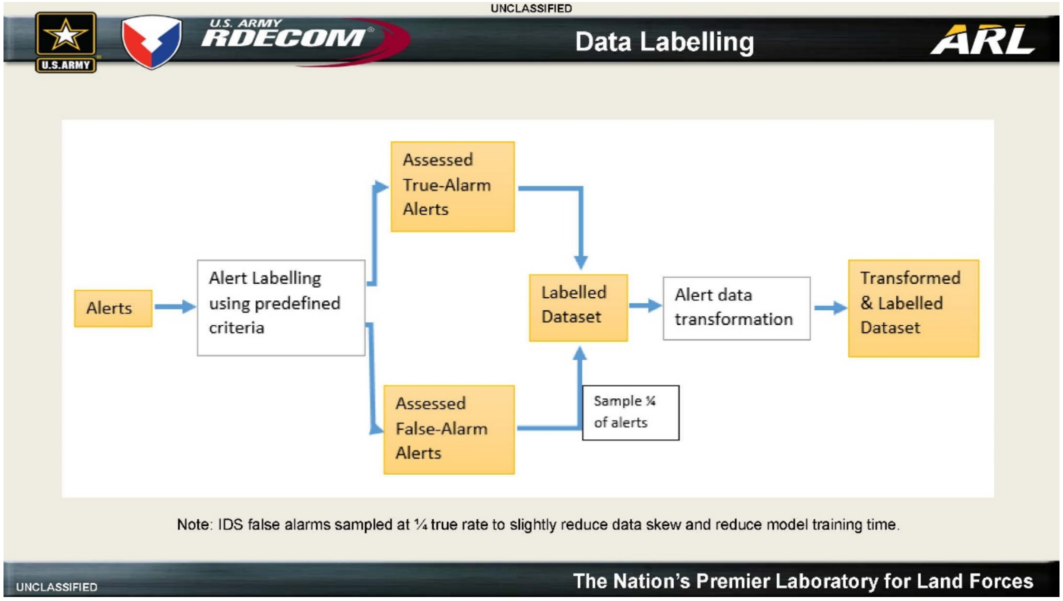
Transformed alert data via a specific methodology to a 23-feature vector

Data Labelling

Define True Positive Alerts:

- **Alert originates from same location (as reported incident)**
- **Alert contains same IP addresses**
- **Alert originates within a +/- 1 hour timeframe of report, OR if alert generating tool matches, alert must originate from within a +/- 24 hour window**

These criteria are necessary because incident reporting process is likely to be human driven – not millisecond exact.



UNCLASSIFIED

Data Labelling

Alerts

Alert Labelling using predefined criteria

Assessed True-Alarm Alerts

Assessed False-Alarm Alerts

Sample ¼ of alerts

Labelled Dataset

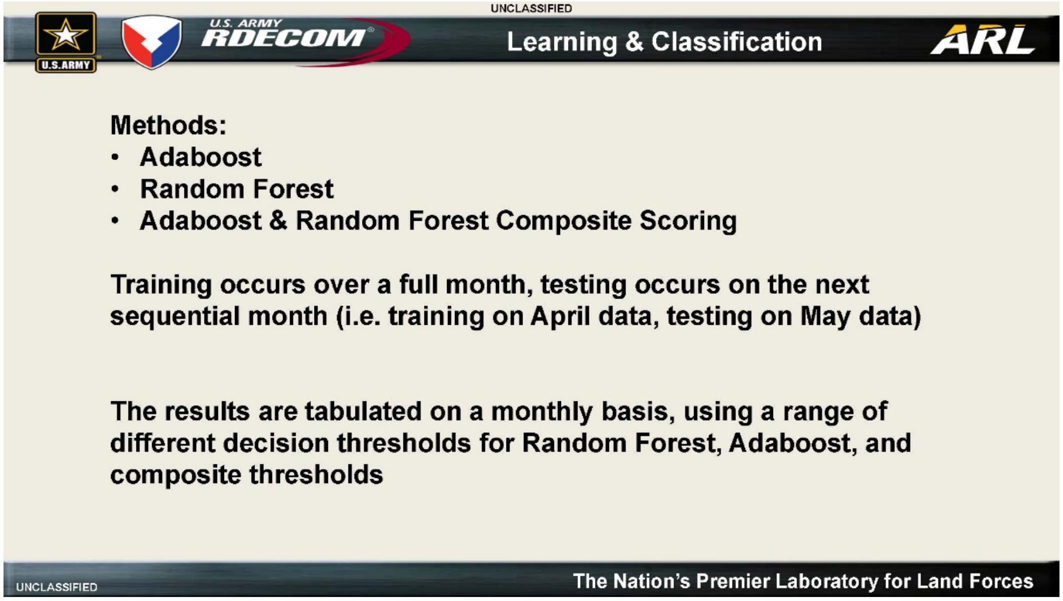
Alert data transformation

Transformed & Labelled Dataset

Note: IDS false alarms sampled at ¼ true rate to slightly reduce data skew and reduce model training time.

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

Learning & Classification

Methods:

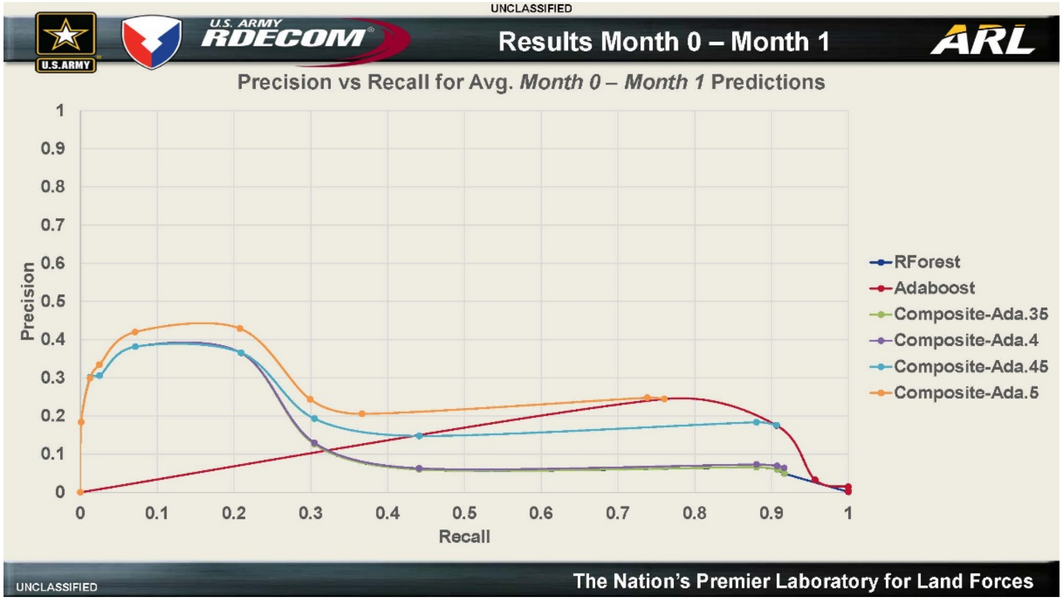
- Adaboost
- Random Forest
- Adaboost & Random Forest Composite Scoring

Training occurs over a full month, testing occurs on the next sequential month (i.e. training on April data, testing on May data)

The results are tabulated on a monthly basis, using a range of different decision thresholds for Random Forest, Adaboost, and composite thresholds

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

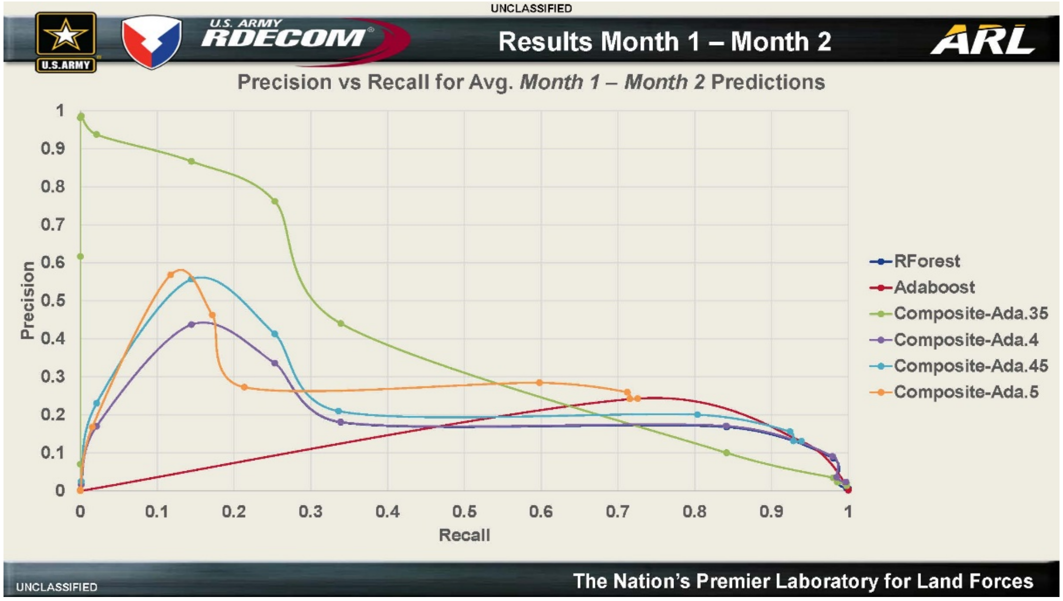
Results Month 0 – Month 1

ARL

Precision vs Recall for Avg. *Month 0 – Month 1* Predictions

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

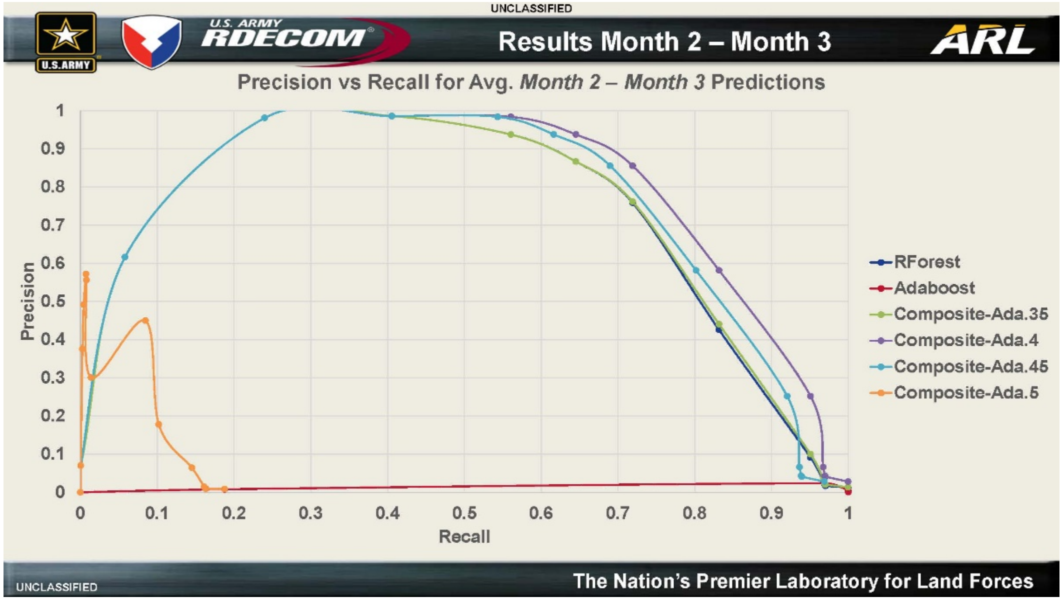
Results Month 1 – Month 2

ARL

Precision vs Recall for Avg. *Month 1 – Month 2* Predictions

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



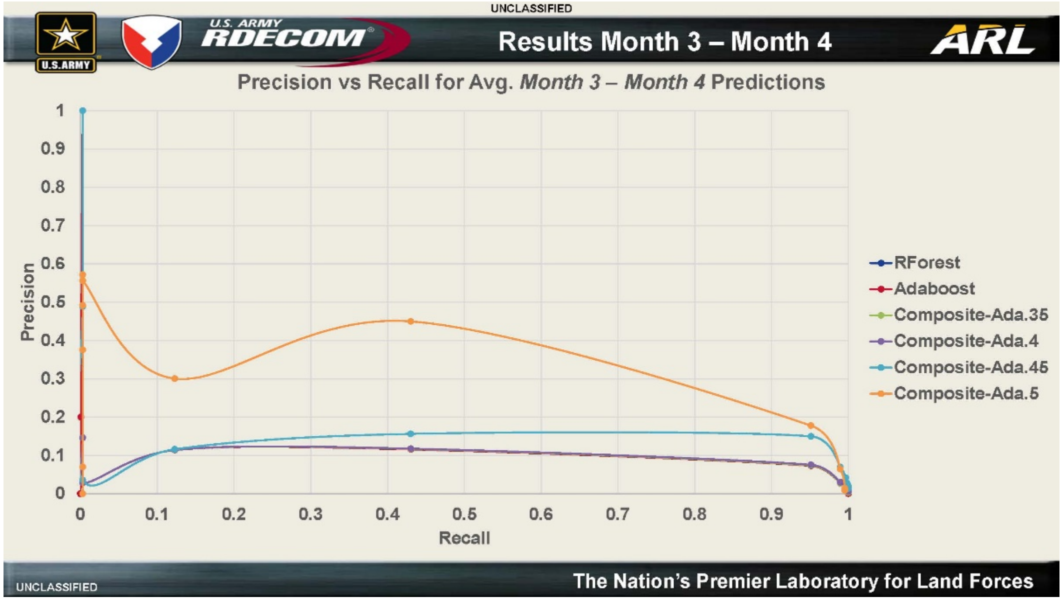
UNCLASSIFIED

Results Month 2 – Month 3

Precision vs Recall for Avg. *Month 2 – Month 3* Predictions

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



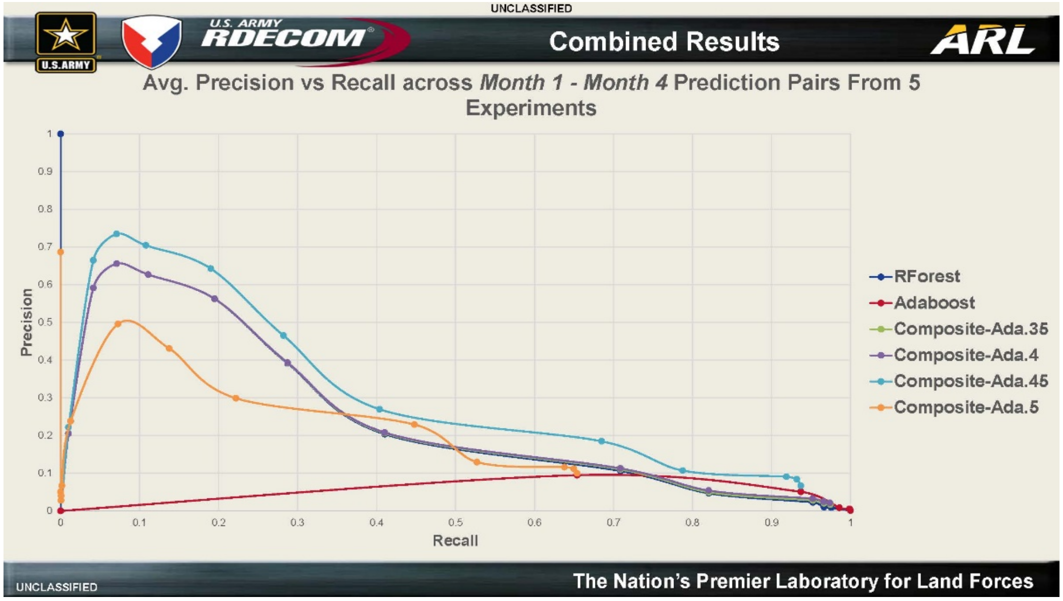
UNCLASSIFIED

Results Month 3 – Month 4

Precision vs Recall for Avg. *Month 3 – Month 4* Predictions

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



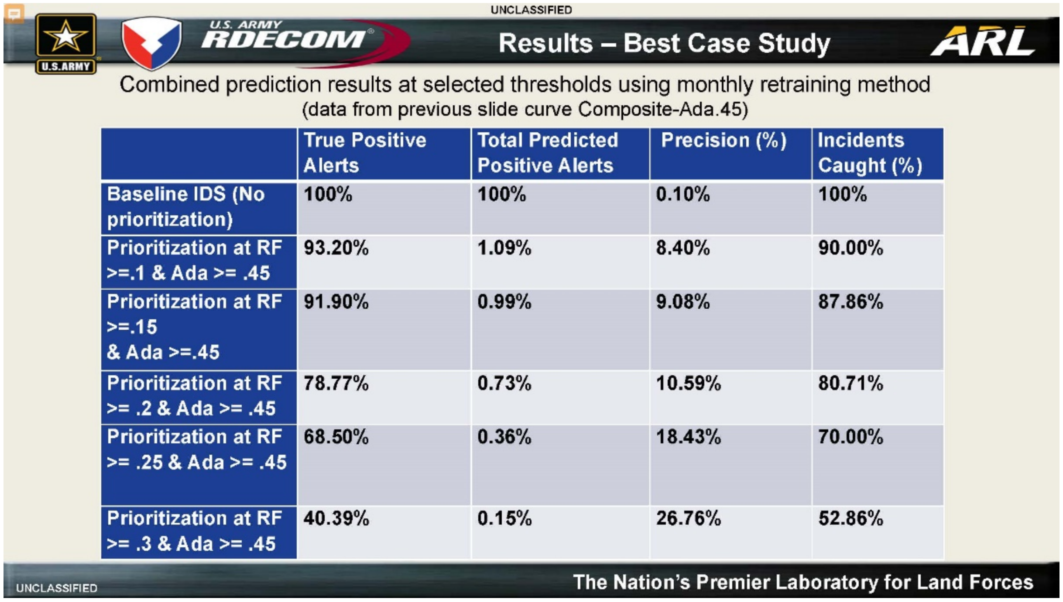
UNCLASSIFIED

Combined Results

Avg. Precision vs Recall across *Month 1 - Month 4* Prediction Pairs From 5 Experiments

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

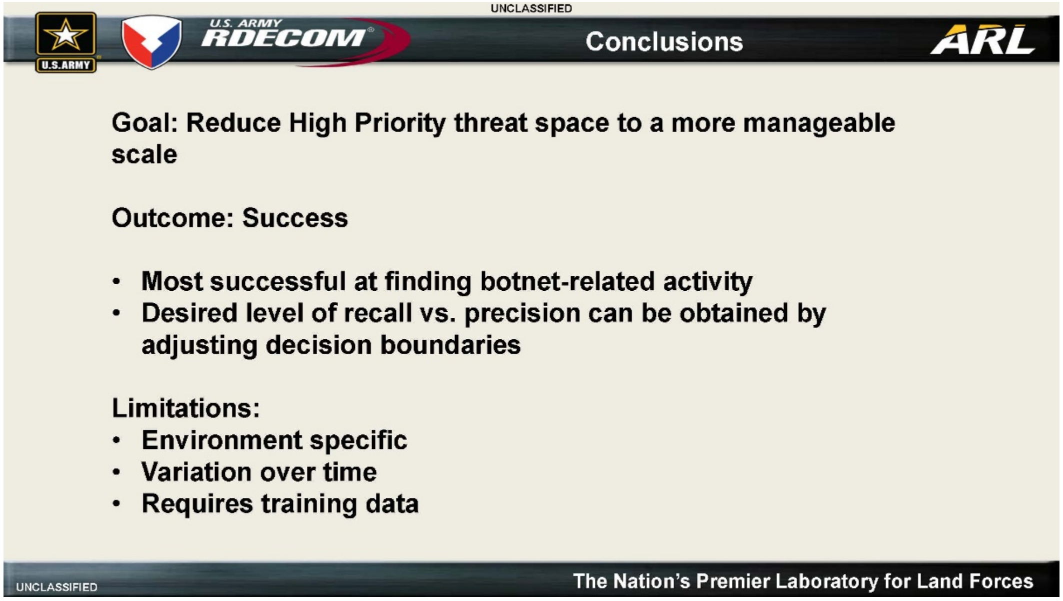
Results – Best Case Study

Combined prediction results at selected thresholds using monthly retraining method
(data from previous slide curve Composite-Ada.45)

	True Positive Alerts	Total Predicted Positive Alerts	Precision (%)	Incidents Caught (%)
Baseline IDS (No prioritization)	100%	100%	0.10%	100%
Prioritization at RF >=.1 & Ada >= .45	93.20%	1.09%	8.40%	90.00%
Prioritization at RF >=.15 & Ada >=.45	91.90%	0.99%	9.08%	87.86%
Prioritization at RF >= .2 & Ada >= .45	78.77%	0.73%	10.59%	80.71%
Prioritization at RF >= .25 & Ada >= .45	68.50%	0.36%	18.43%	70.00%
Prioritization at RF >= .3 & Ada >= .45	40.39%	0.15%	26.76%	52.86%

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



U.S. ARMY
U.S. ARMY RDECOM
ARL

Conclusions

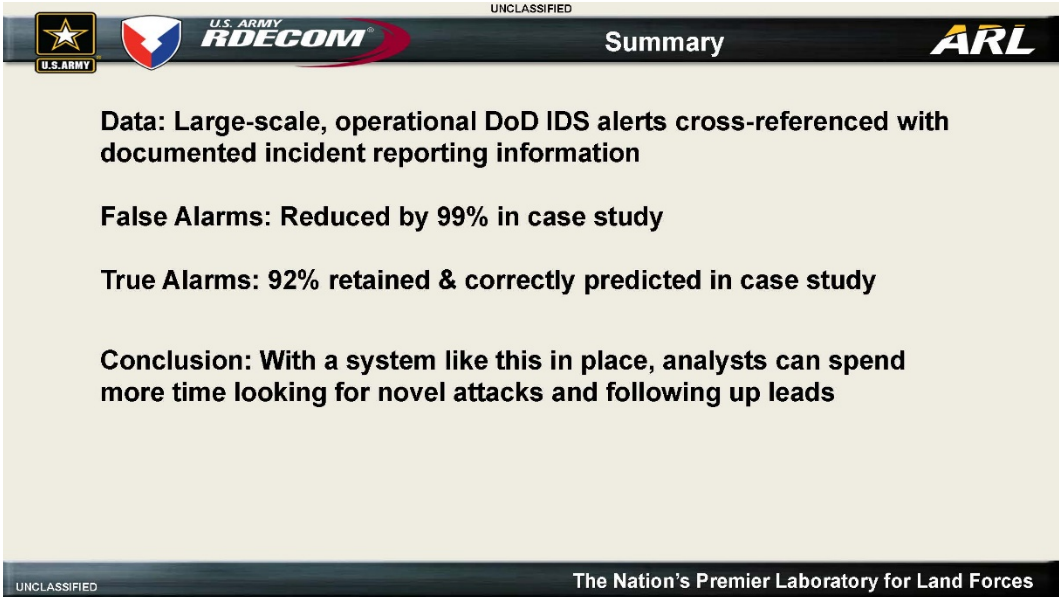
Goal: Reduce High Priority threat space to a more manageable scale

Outcome: Success

- Most successful at finding botnet-related activity
- Desired level of recall vs. precision can be obtained by adjusting decision boundaries

Limitations:

- Environment specific
- Variation over time
- Requires training data



U.S. ARMY
U.S. ARMY RDECOM
ARL

Summary

Data: Large-scale, operational DoD IDS alerts cross-referenced with documented incident reporting information

False Alarms: Reduced by 99% in case study

True Alarms: 92% retained & correctly predicted in case study

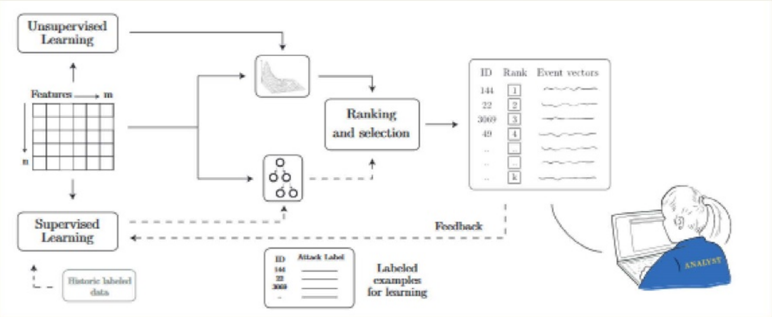
Conclusion: With a system like this in place, analysts can spend more time looking for novel attacks and following up leads

Future/Improvements

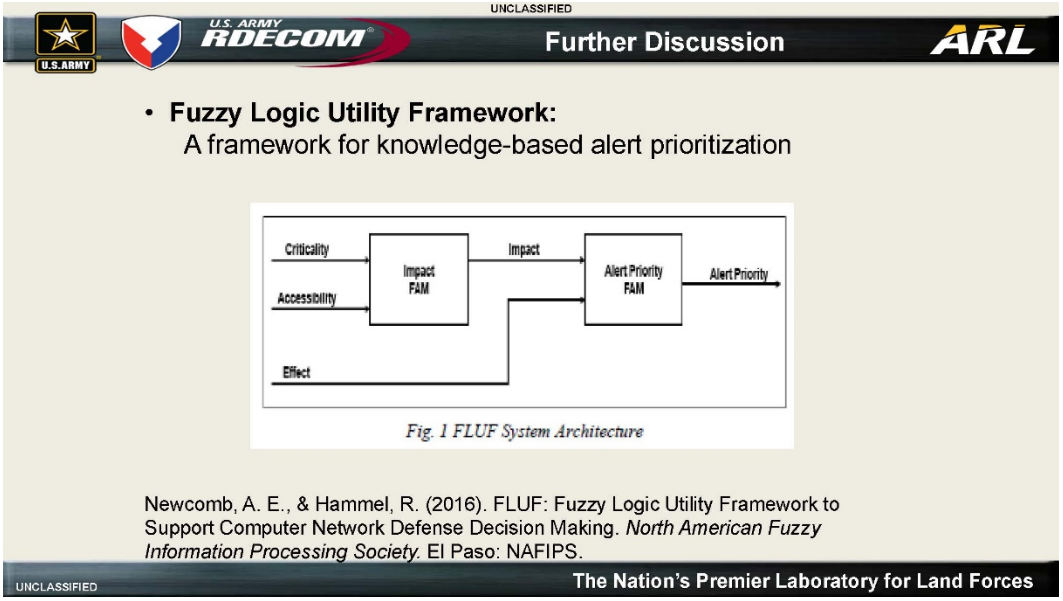
- **Investigation of labelling errors – reassessing false positives/false negatives for relabeling**
- **Data -> Features transformation methodology improvements**
- **Training vs. Testing Period (Timespan, sliding window, etc.)**
- **Incorporation of a pre-established alert priority feature**
- **Better documentation of alerts within incident reporting framework to allow for more precise correlation**

Further Discussion

- **AI^2: Training a big data machine to defend**
  A study by MIT CSAIL Laboratory and PatternEx



Veeramachaneni, K., Arnaldo, I., Korrapati, V., Bassias, C., & Li, K. (2016). AI2: Training a big data machine to defend. *IEEE 2nd International Conference on Big Data Security on Cloud.* IEEE.



UNCLASSIFIED

U.S. ARMY RDECOM

Further Discussion

ARL

- **Fuzzy Logic Utility Framework:**
  A framework for knowledge-based alert prioritization

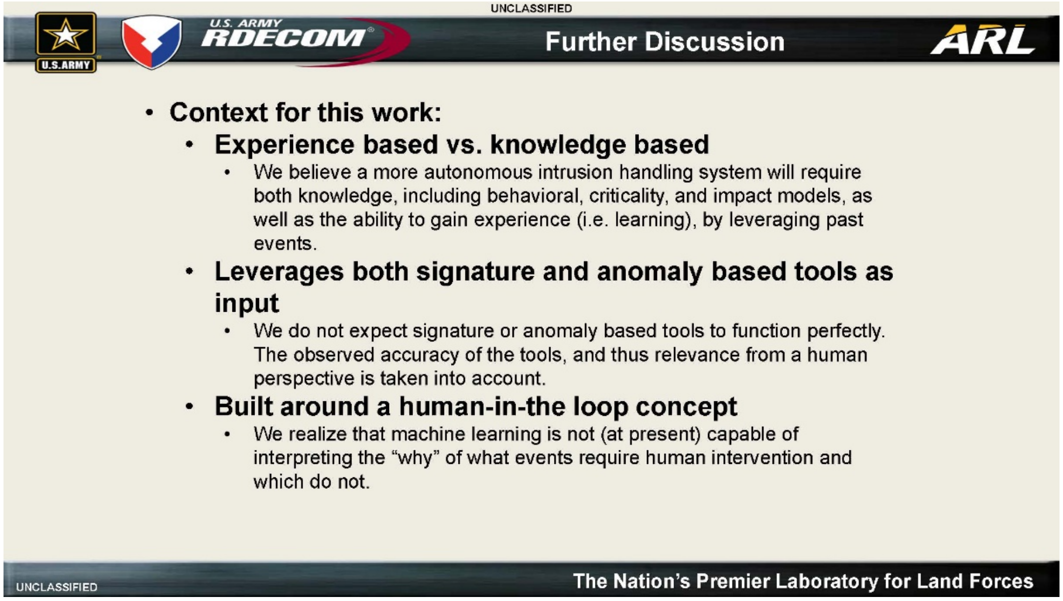
Criticality
Accessibility
Impact FAM
Impact
Alert Priority FAM
Alert Priority
Effect

Fig. 1 FLUF System Architecture

Newcomb, A. E., & Hammel, R. (2016). FLUF: Fuzzy Logic Utility Framework to Support Computer Network Defense Decision Making. *North American Fuzzy Information Processing Society.* El Paso: NAFIPS.

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

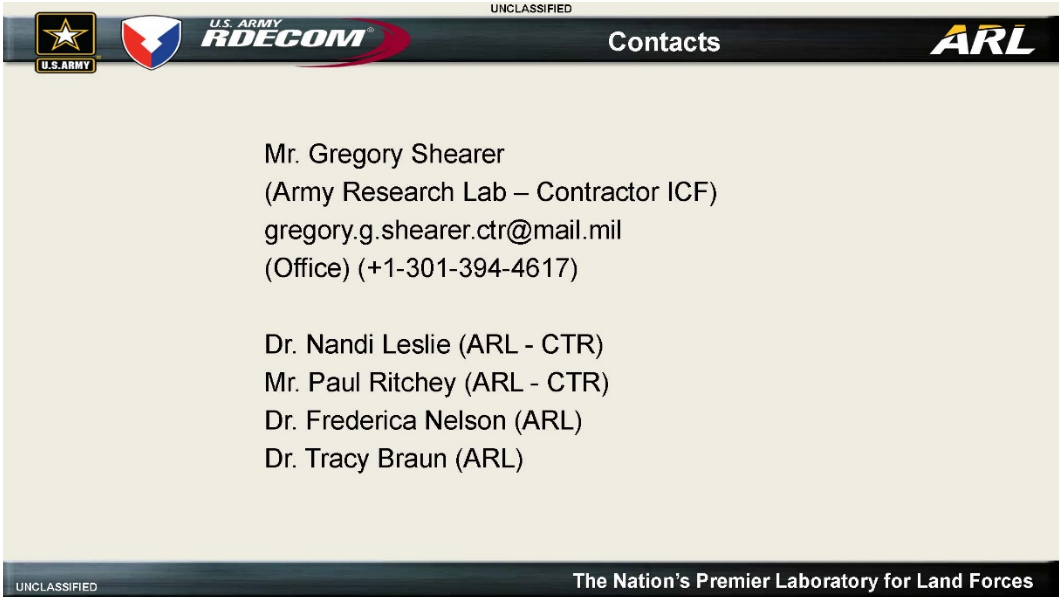
Further Discussion

ARL

- **Context for this work:**
  - **Experience based vs. knowledge based**
    - We believe a more autonomous intrusion handling system will require both knowledge, including behavioral, criticality, and impact models, as well as the ability to gain experience (i.e. learning), by leveraging past events.
  - **Leverages both signature and anomaly based tools as input**
    - We do not expect signature or anomaly based tools to function perfectly. The observed accuracy of the tools, and thus relevance from a human perspective is taken into account.
  - **Built around a human-in-the loop concept**
    - We realize that machine learning is not (at present) capable of interpreting the "why" of what events require human intervention and which do not.

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

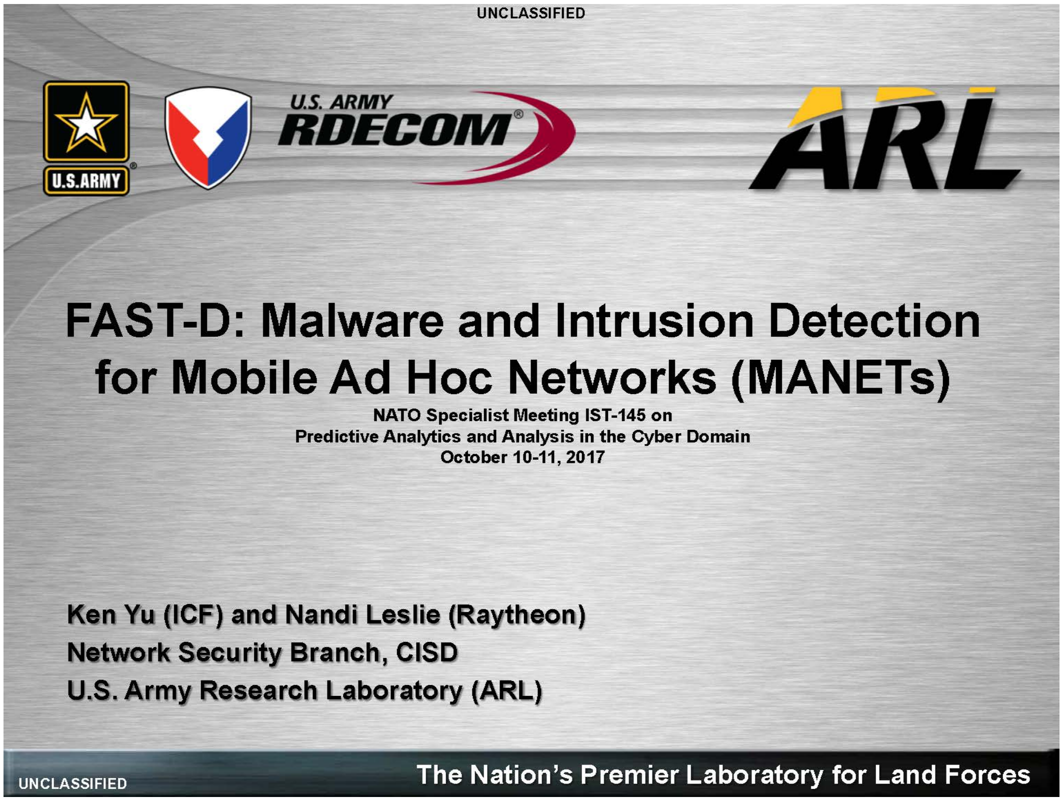
Contacts

Mr. Gregory Shearer
(Army Research Lab – Contractor ICF)
gregory.g.shearer.ctr@mail.mil
(Office) (+1-301-394-4617)

Dr. Nandi Leslie (ARL - CTR)
Mr. Paul Ritchey (ARL - CTR)
Dr. Frederica Nelson (ARL)
Dr. Tracy Braun (ARL)

UNCLASSIFIED

The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

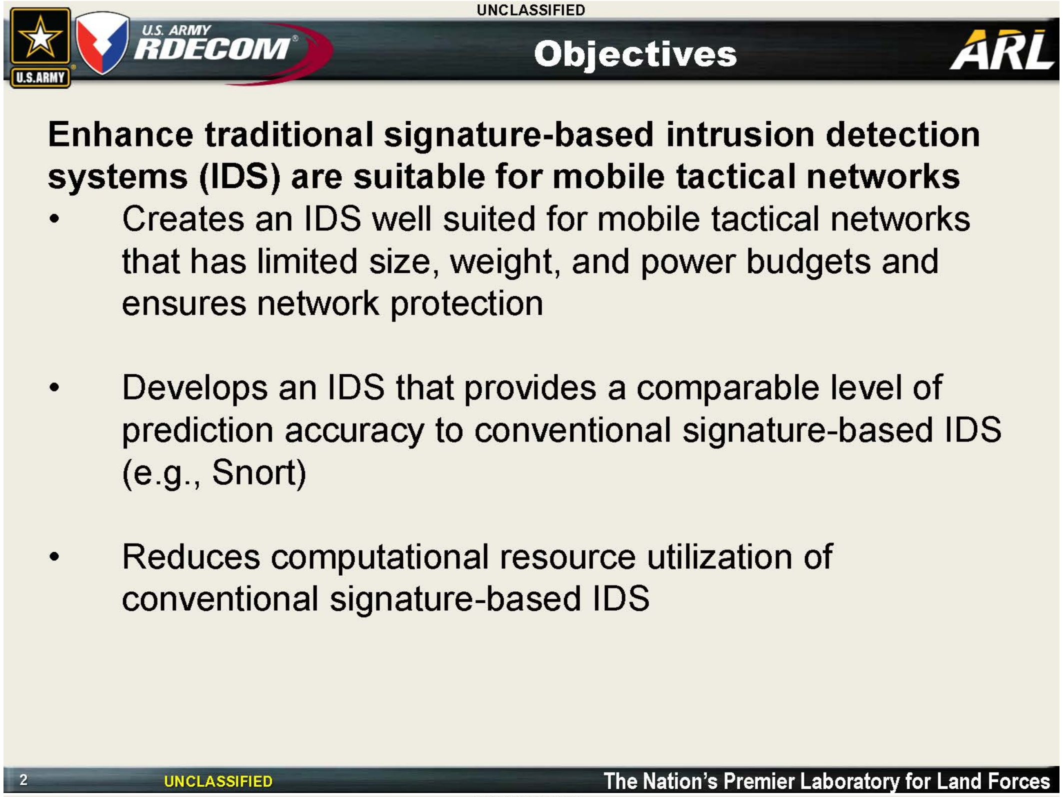
FAST-D: Malware and Intrusion Detection for Mobile Ad Hoc Networks (MANETs)

NATO Specialist Meeting IST-145 on
Predictive Analytics and Analysis in the Cyber Domain
October 10-11, 2017

Ken Yu (ICF) and Nandi Leslie (Raytheon)
Network Security Branch, CISD
U.S. Army Research Laboratory (ARL)

UNCLASSIFIED

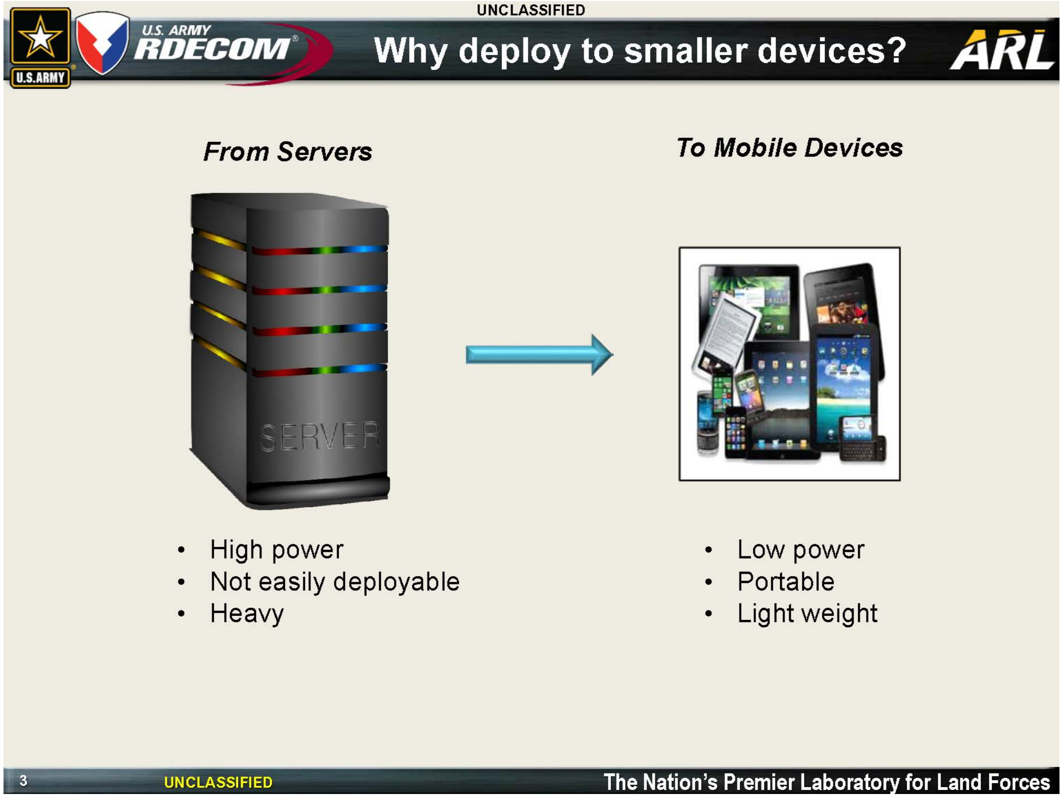
The Nation's Premier Laboratory for Land Forces



Objectives

Enhance traditional signature-based intrusion detection systems (IDS) are suitable for mobile tactical networks

- Creates an IDS well suited for mobile tactical networks that has limited size, weight, and power budgets and ensures network protection
- Develops an IDS that provides a comparable level of prediction accuracy to conventional signature-based IDS (e.g., Snort)
- Reduces computational resource utilization of conventional signature-based IDS



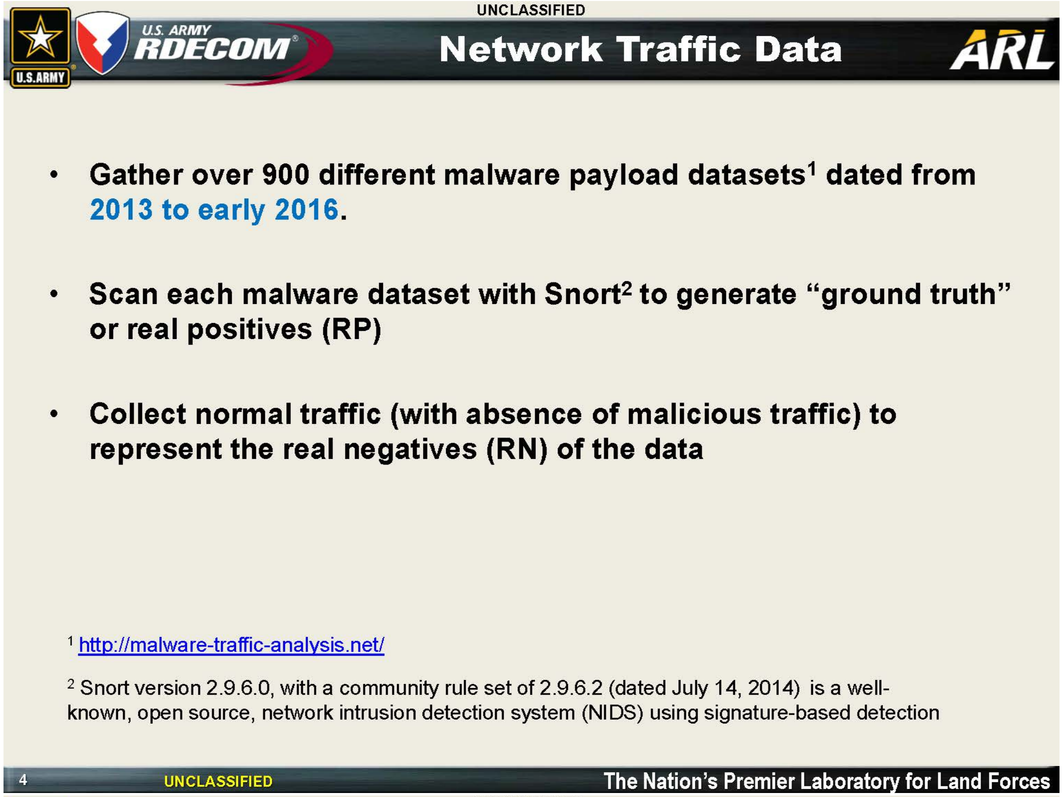
Why deploy to smaller devices?

From Servers

- High power
- Not easily deployable
- Heavy

To Mobile Devices

- Low power
- Portable
- Light weight

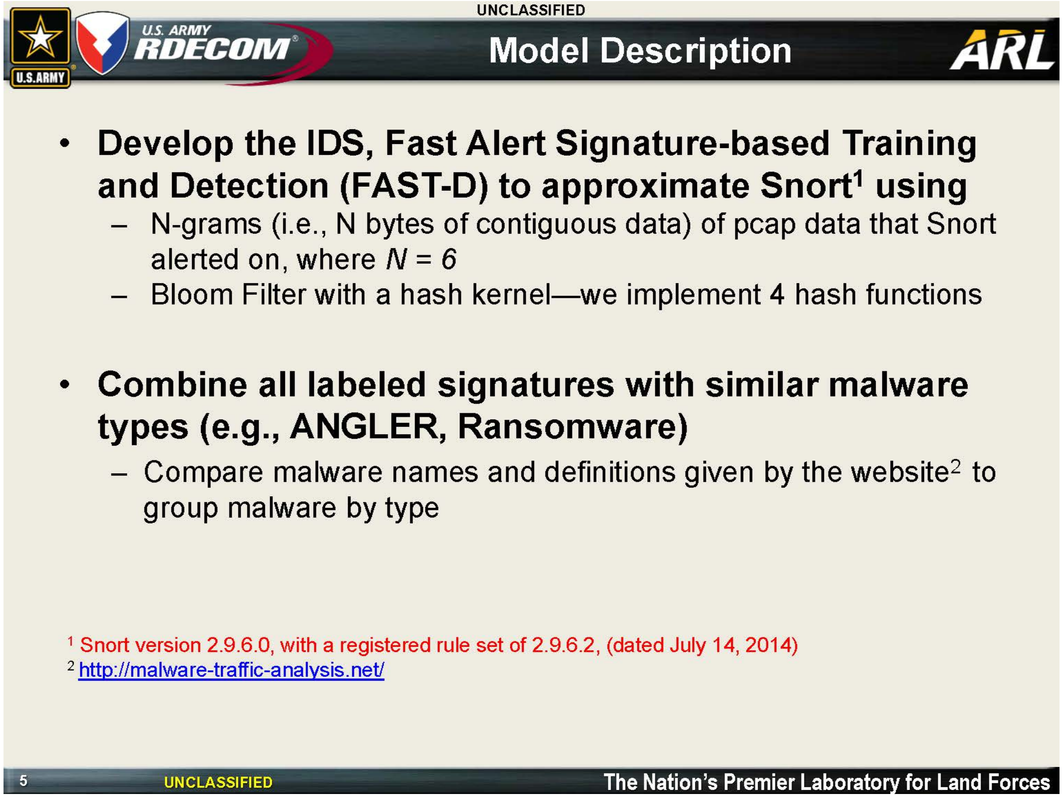


Network Traffic Data

- Gather over 900 different malware payload datasets[1] dated from 2013 to early 2016.
- Scan each malware dataset with Snort[2] to generate “ground truth” or real positives (RP)
- Collect normal traffic (with absence of malicious traffic) to represent the real negatives (RN) of the data

[1] http://malware-traffic-analysis.net/

[2] Snort version 2.9.6.0, with a community rule set of 2.9.6.2 (dated July 14, 2014) is a well-known, open source, network intrusion detection system (NIDS) using signature-based detection

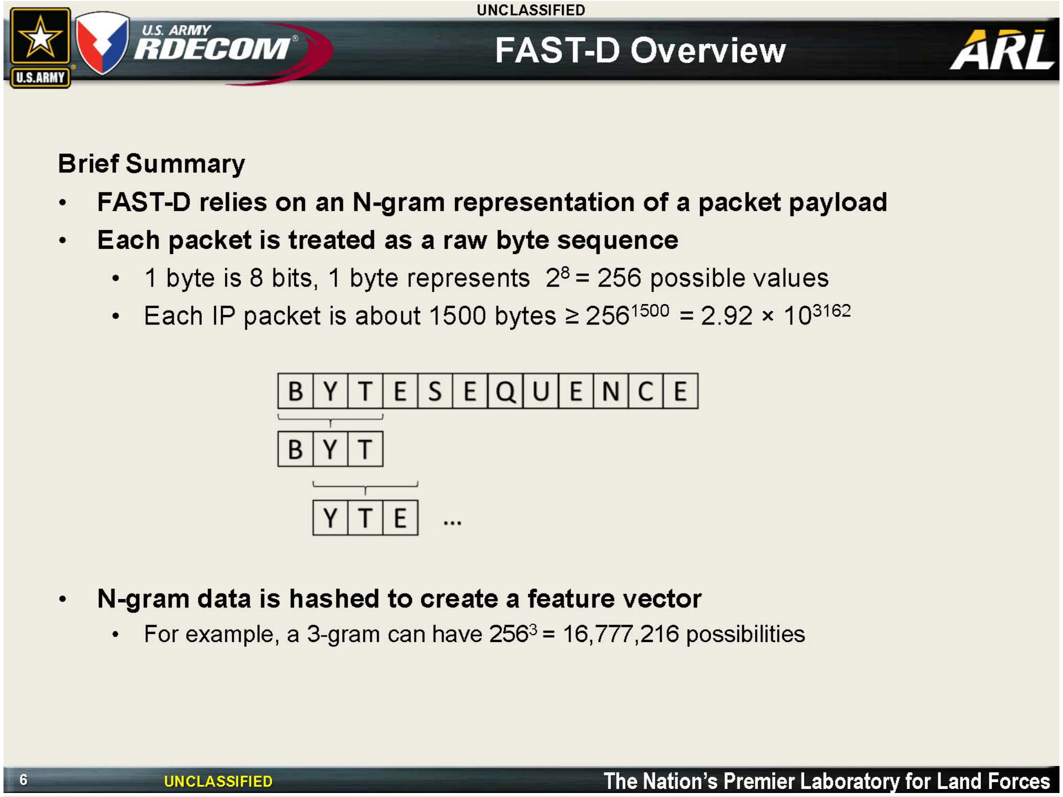


Model Description

- Develop the IDS, Fast Alert Signature-based Training and Detection (FAST-D) to approximate Snort[1] using
  - N-grams (i.e., N bytes of contiguous data) of pcap data that Snort alerted on, where *N = 6*
  - Bloom Filter with a hash kernel—we implement 4 hash functions
- Combine all labeled signatures with similar malware types (e.g., ANGLER, Ransomware)
  - Compare malware names and definitions given by the website[2] to group malware by type

[1] Snort version 2.9.6.0, with a registered rule set of 2.9.6.2, (dated July 14, 2014)

[2] http://malware-traffic-analysis.net/



FAST-D Overview

Brief Summary

- **FAST-D relies on an N-gram representation of a packet payload**
- **Each packet is treated as a raw byte sequence**
  - 1 byte is 8 bits, 1 byte represents $2^8 = 256$ possible values
  - Each IP packet is about 1500 bytes $\geq 256^{1500} = 2.92 \times 10^{3162}$

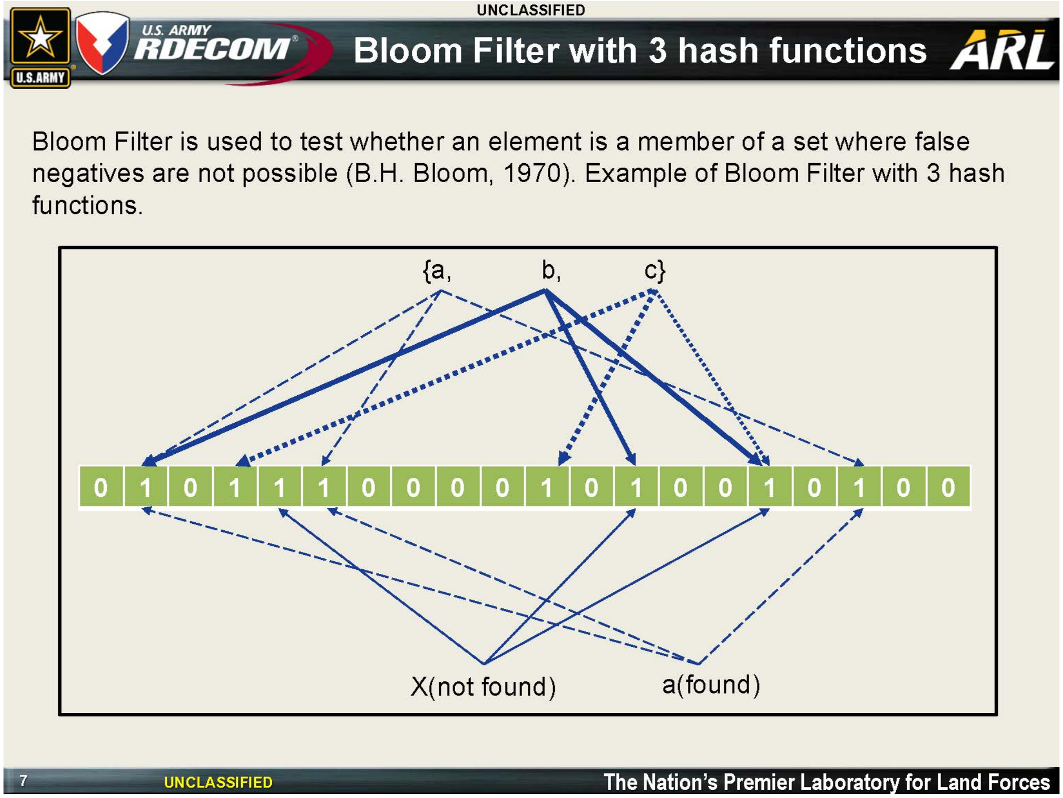
B	Y	T	E	S	E	Q	U	E	N	C	E

B	Y	T

Y	T	E

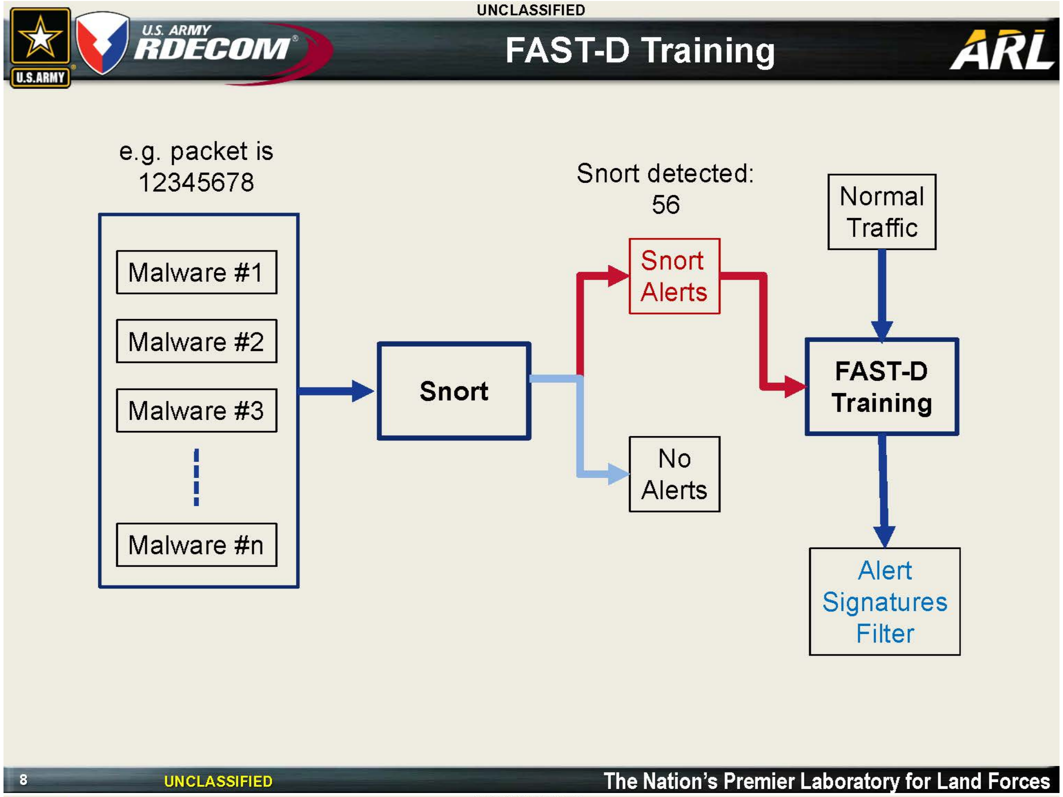
…

- **N-gram data is hashed to create a feature vector**
  - For example, a 3-gram can have $256^3 = 16{,}777{,}216$ possibilities



Bloom Filter with 3 hash functions

Bloom Filter is used to test whether an element is a member of a set where false negatives are not possible (B.H. Bloom, 1970). Example of Bloom Filter with 3 hash functions.



UNCLASSIFIED

U.S. ARMY RDECOM

FAST-D Training

ARL

e.g. packet is 12345678

Malware #1

Malware #2

Malware #3

Malware #n

Snort

Snort detected: 56

Snort Alerts

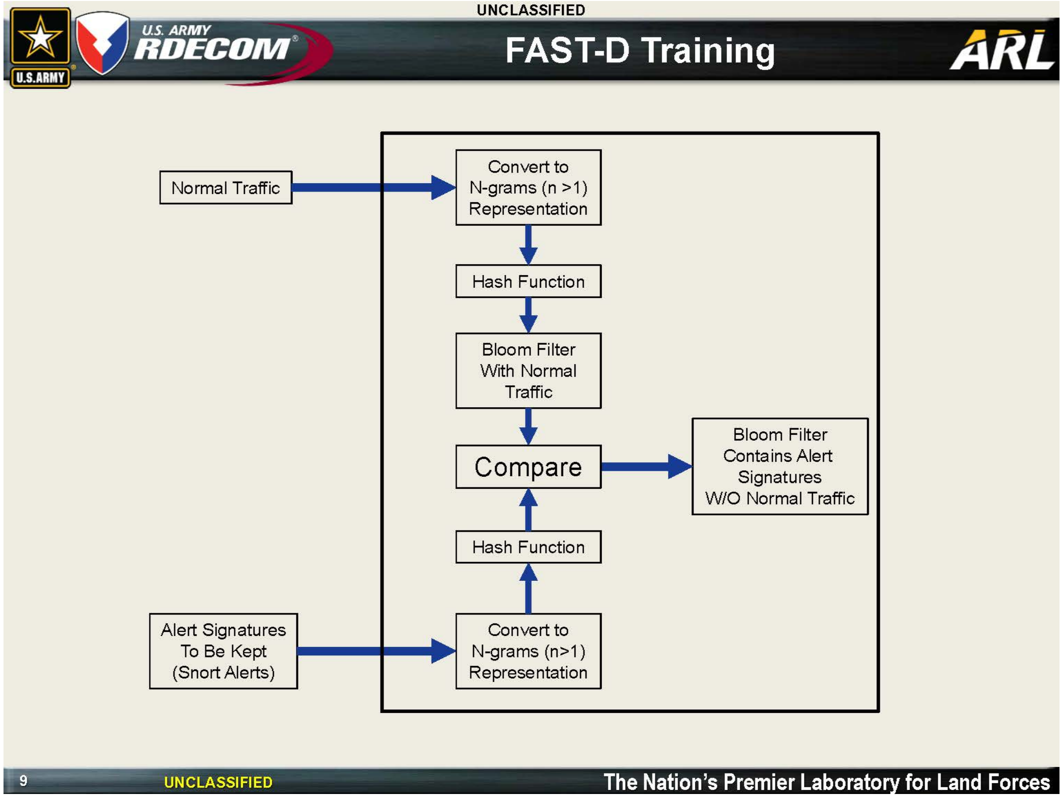
No Alerts

Normal Traffic

FAST-D Training

Alert Signatures Filter

8 UNCLASSIFIED The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

FAST-D Training

ARL

Normal Traffic

Convert to N-grams (n >1) Representation

Hash Function

Bloom Filter With Normal Traffic

Compare

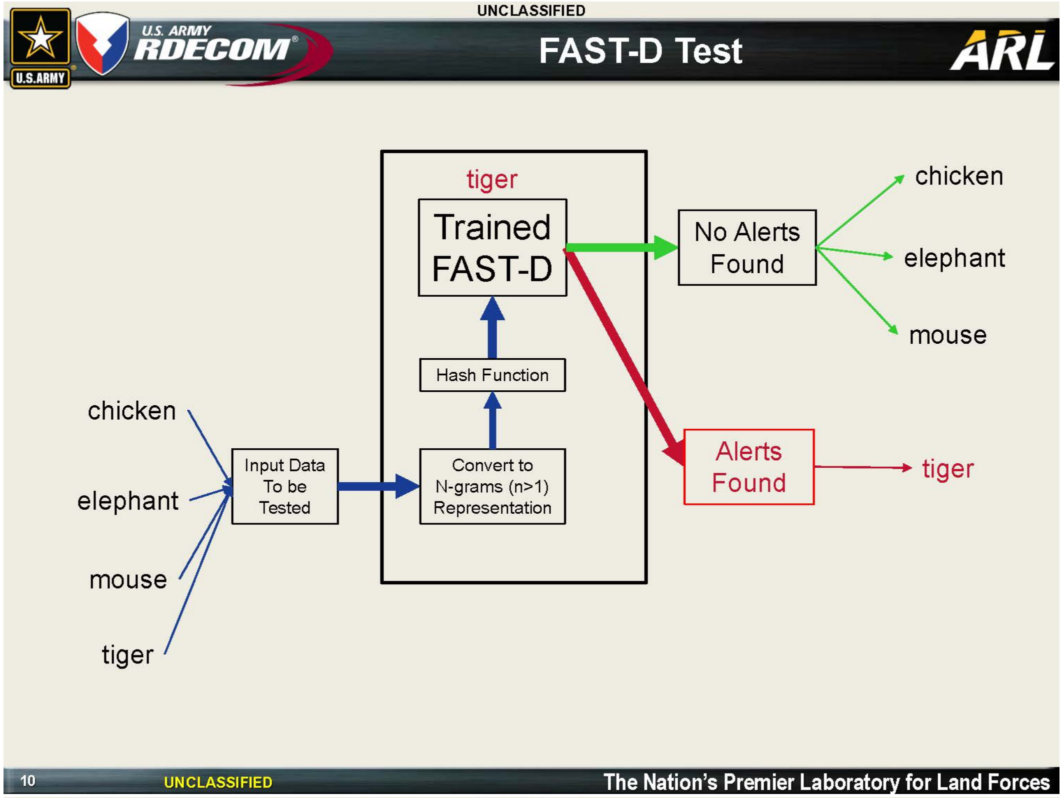
Bloom Filter Contains Alert Signatures W/O Normal Traffic

Hash Function

Convert to N-grams (n>1) Representation

Alert Signatures To Be Kept (Snort Alerts)

9 UNCLASSIFIED The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

U.S. ARMY RDECOM

FAST-D Test

ARL

chicken
elephant
mouse
tiger

Input Data To be Tested

Convert to N-grams (n>1) Representation

Hash Function

Trained FAST-D

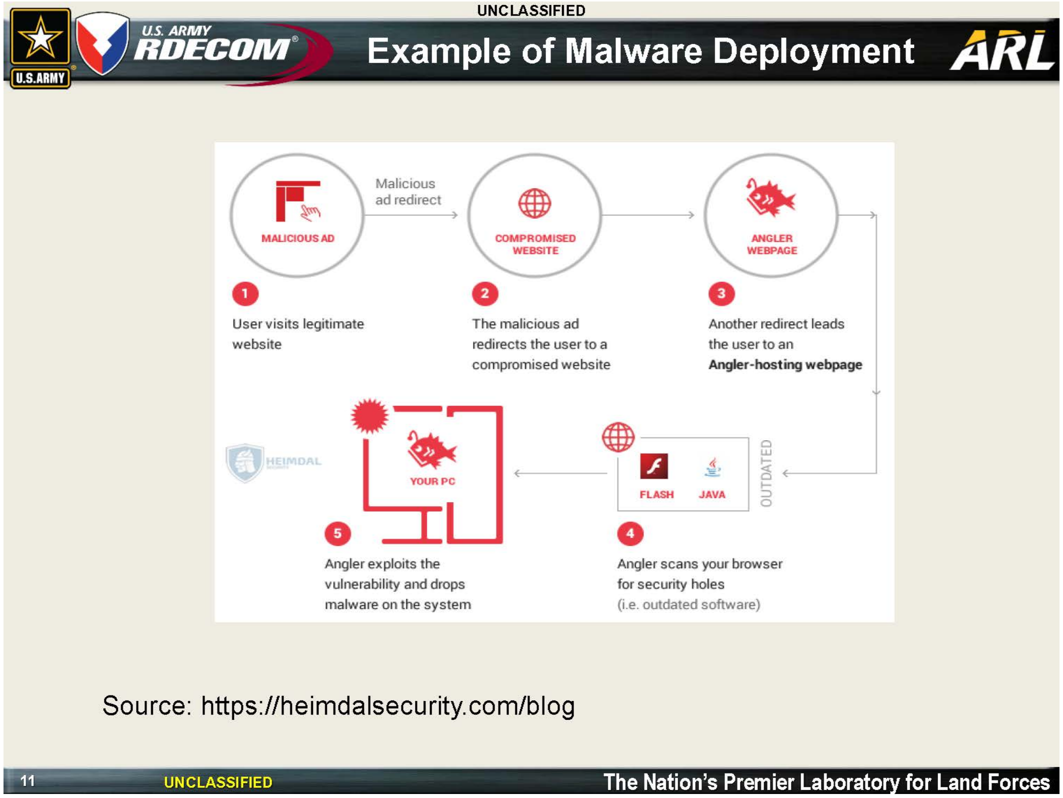
tiger

No Alerts Found

chicken
elephant
mouse

Alerts Found

tiger



UNCLASSIFIED

U.S. ARMY RDECOM

Example of Malware Deployment

ARL

MALICIOUS AD

Malicious ad redirect

COMPROMISED WEBSITE

ANGLER WEBPAGE

1 User visits legitimate website

2 The malicious ad redirects the user to a compromised website

3 Another redirect leads the user to an **Angler-hosting webpage**

HEIMDAL

YOUR PC

FLASH

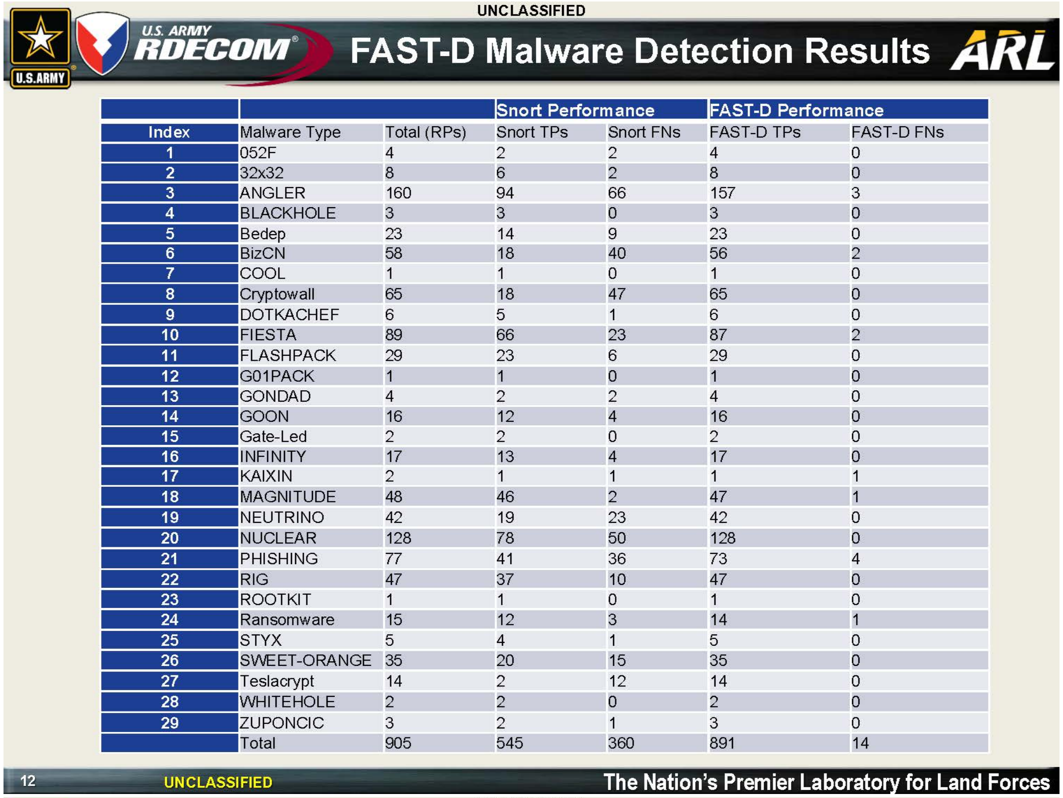
JAVA

OUTDATED

5 Angler exploits the vulnerability and drops malware on the system

4 Angler scans your browser for security holes (i.e. outdated software)

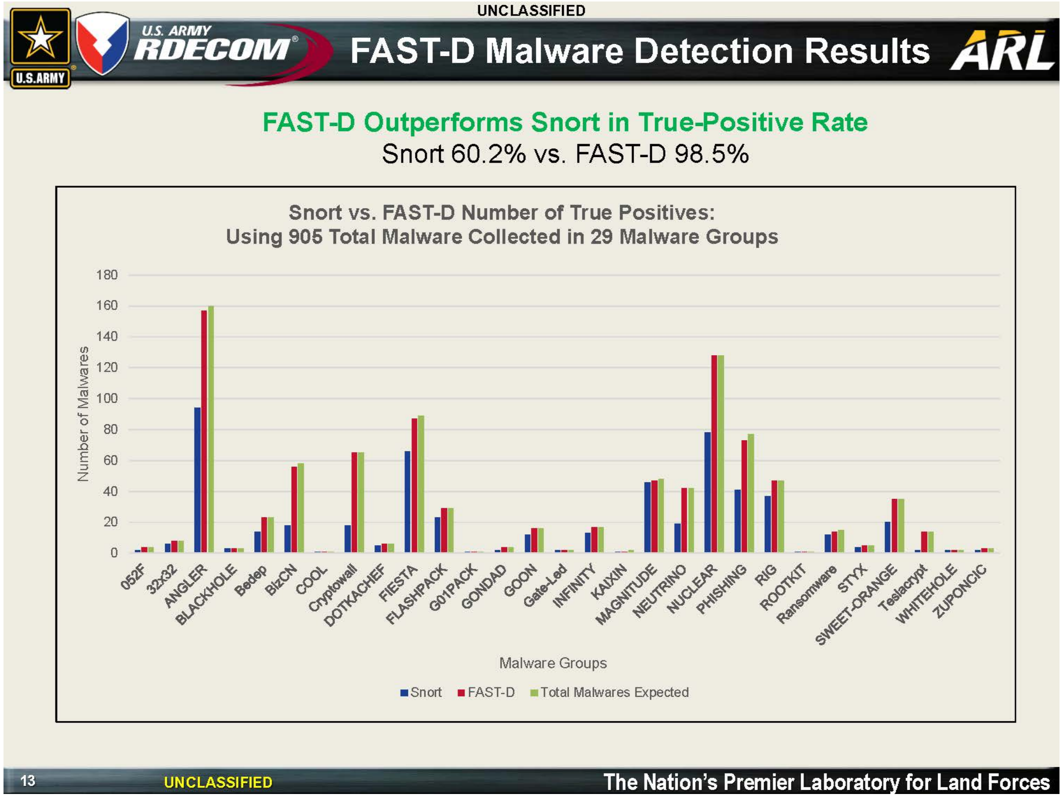
Source: https://heimdalsecurity.com/blog



U.S. ARMY RDECOM® U.S.ARMY ARL

FAST-D Malware Detection Results

			Snort Performance		FAST-D Performance	
Index	Malware Type	Total (RPs)	Snort TPs	Snort FNs	FAST-D TPs	FAST-D FNs
1	052F	4	2	2	4	0
2	32x32	8	6	2	8	0
3	ANGLER	160	94	66	157	3
4	BLACKHOLE	3	3	0	3	0
5	Bedep	23	14	9	23	0
6	BizCN	58	18	40	56	2
7	COOL	1	1	0	1	0
8	Cryptowall	65	18	47	65	0
9	DOTKACHEF	6	5	1	6	0
10	FIESTA	89	66	23	87	2
11	FLASHPACK	29	23	6	29	0
12	G01PACK	1	1	0	1	0
13	GONDAD	4	2	2	4	0
14	GOON	16	12	4	16	0
15	Gate-Led	2	2	0	2	0
16	INFINITY	17	13	4	17	0
17	KAIXIN	2	1	1	1	1
18	MAGNITUDE	48	46	2	47	1
19	NEUTRINO	42	19	23	42	0
20	NUCLEAR	128	78	50	128	0
21	PHISHING	77	41	36	73	4
22	RIG	47	37	10	47	0
23	ROOTKIT	1	1	0	1	0
24	Ransomware	15	12	3	14	1
25	STYX	5	4	1	5	0
26	SWEET-ORANGE	35	20	15	35	0
27	Teslacrypt	14	2	12	14	0
28	WHITEHOLE	2	2	0	2	0
29	ZUPONCIC	3	2	1	3	0
	Total	905	545	360	891	14



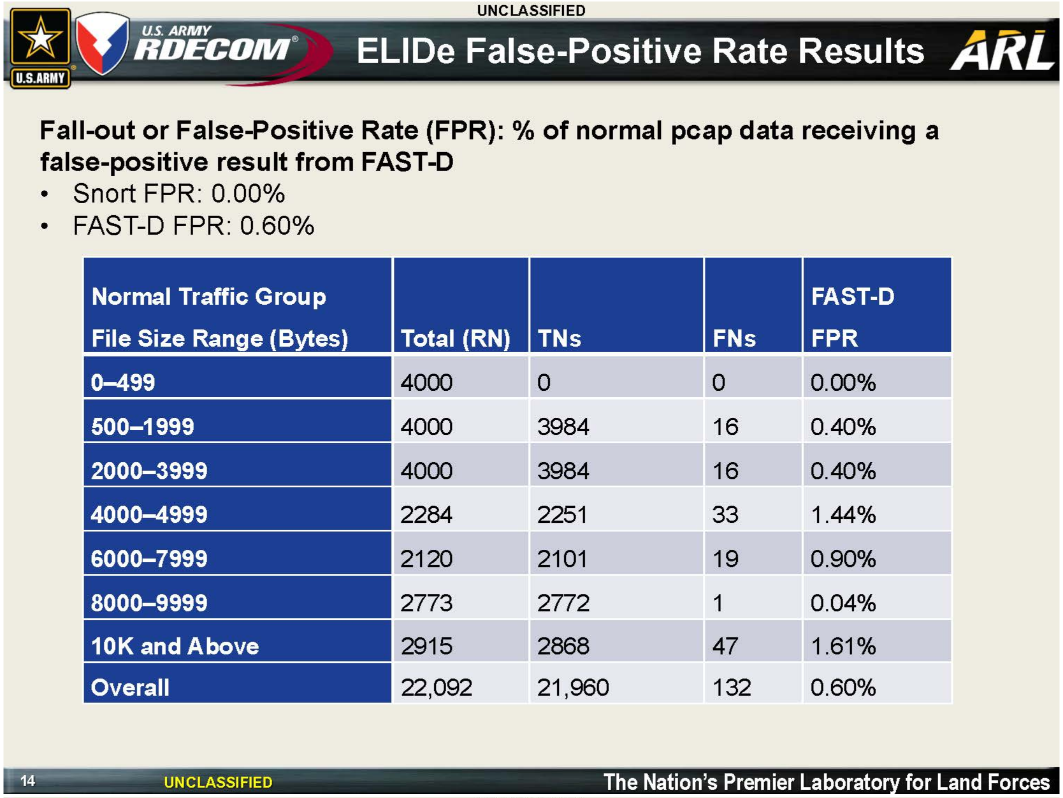
U.S. ARMY RDECOM® U.S.ARMY ARL

FAST-D Malware Detection Results

FAST-D Outperforms Snort in True-Positive Rate

Snort 60.2% vs. FAST-D 98.5%

Snort vs. FAST-D Number of True Positives: Using 905 Total Malware Collected in 29 Malware Groups

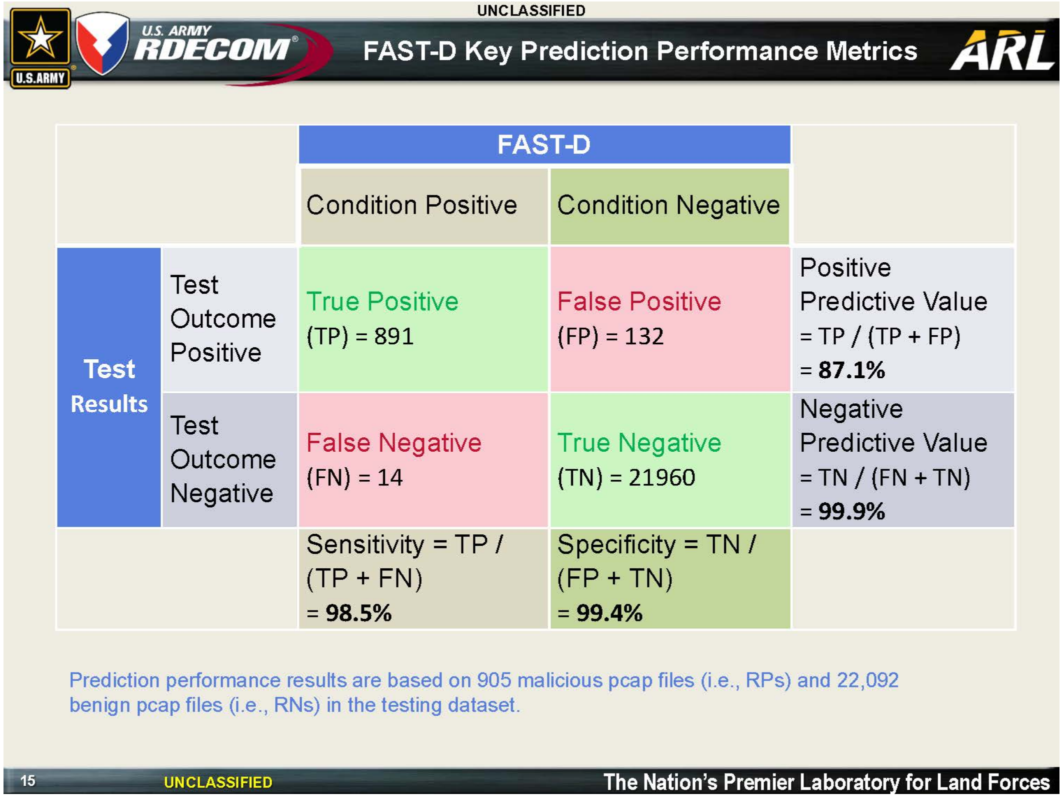


ELIDe False-Positive Rate Results

Fall-out or False-Positive Rate (FPR): % of normal pcap data receiving a false-positive result from FAST-D

- Snort FPR: 0.00%
- FAST-D FPR: 0.60%

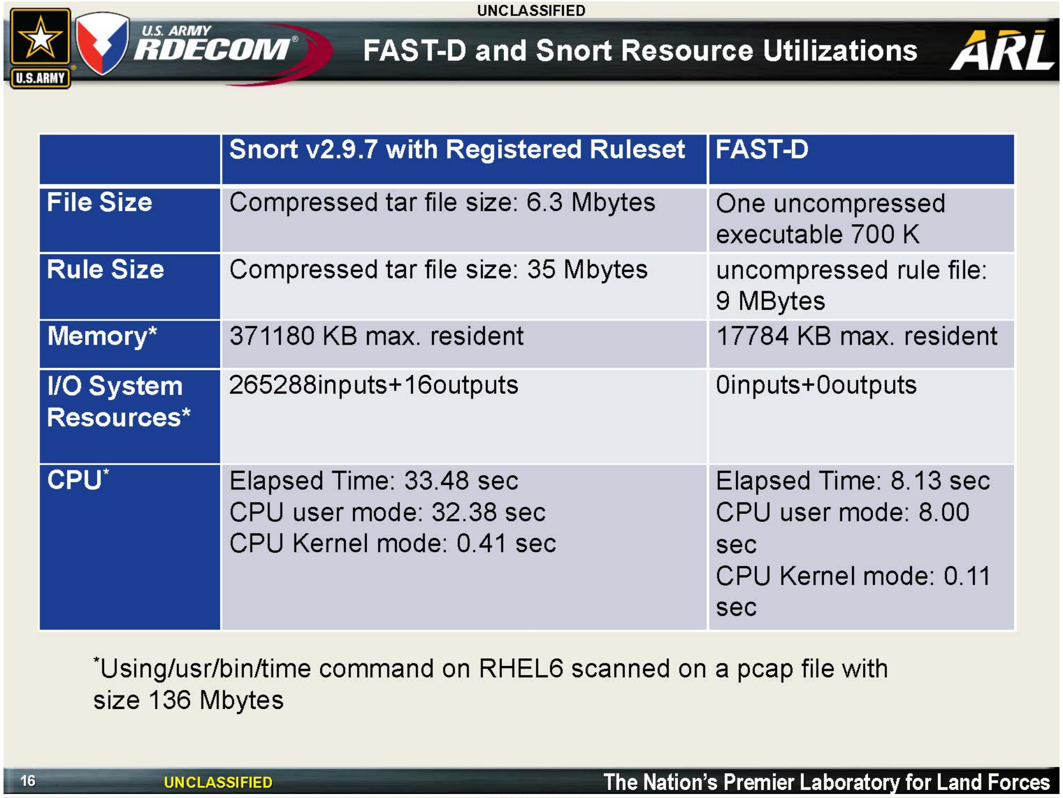
Normal Traffic Group File Size Range (Bytes)	Total (RN)	TNs	FNs	FAST-D FPR
0–499	4000	0	0	0.00%
500–1999	4000	3984	16	0.40%
2000–3999	4000	3984	16	0.40%
4000–4999	2284	2251	33	1.44%
6000–7999	2120	2101	19	0.90%
8000–9999	2773	2772	1	0.04%
10K and Above	2915	2868	47	1.61%
Overall	22,092	21,960	132	0.60%



FAST-D Key Prediction Performance Metrics

		FAST-D		
		Condition Positive	Condition Negative	
Test Results	Test Outcome Positive	True Positive (TP) = 891	False Positive (FP) = 132	Positive Predictive Value = TP / (TP + FP) = **87.1%**
	Test Outcome Negative	False Negative (FN) = 14	True Negative (TN) = 21960	Negative Predictive Value = TN / (FN + TN) = **99.9%**
		Sensitivity = TP / (TP + FN) = **98.5%**	Specificity = TN / (FP + TN) = **99.4%**	

Prediction performance results are based on 905 malicious pcap files (i.e., RPs) and 22,092 benign pcap files (i.e., RNs) in the testing dataset.



UNCLASSIFIED

U.S. ARMY RDECOM

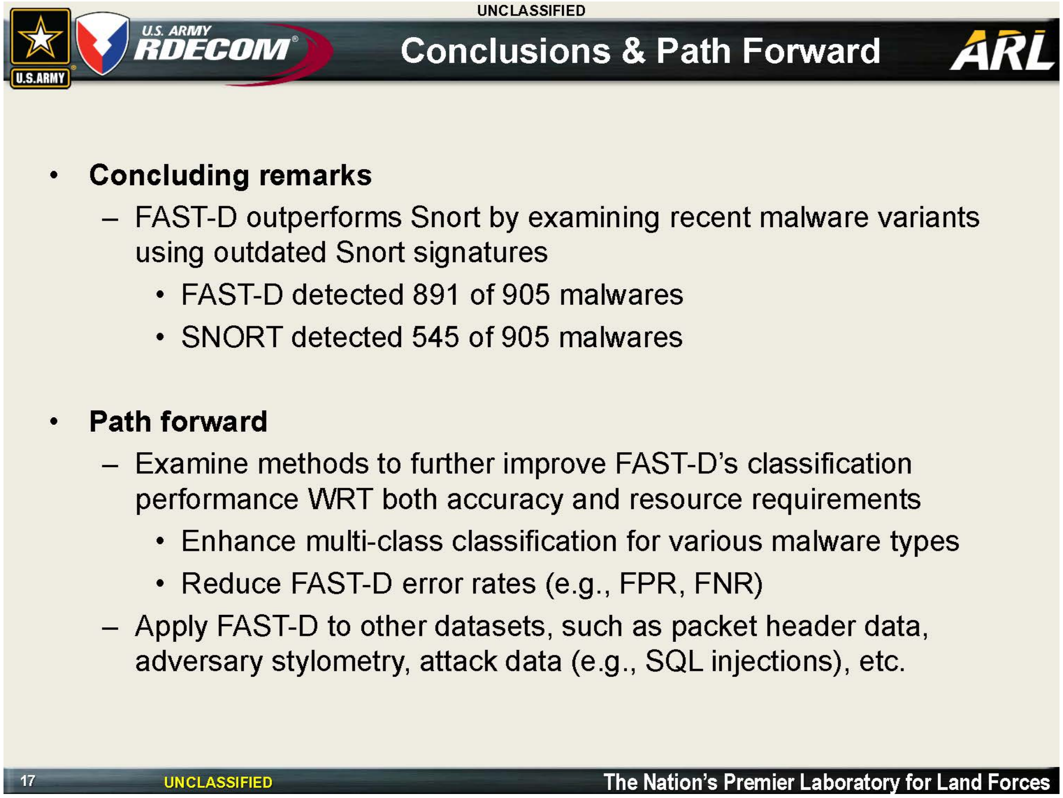
FAST-D and Snort Resource Utilizations

ARL

	Snort v2.9.7 with Registered Ruleset	FAST-D
File Size	Compressed tar file size: 6.3 Mbytes	One uncompressed executable 700 K
Rule Size	Compressed tar file size: 35 Mbytes	uncompressed rule file: 9 MBytes
Memory*	371180 KB max. resident	17784 KB max. resident
I/O System Resources*	265288inputs+16outputs	0inputs+0outputs
CPU*	Elapsed Time: 33.48 sec CPU user mode: 32.38 sec CPU Kernel mode: 0.41 sec	Elapsed Time: 8.13 sec CPU user mode: 8.00 sec CPU Kernel mode: 0.11 sec

*Using/usr/bin/time command on RHEL6 scanned on a pcap file with size 136 Mbytes

16 UNCLASSIFIED The Nation's Premier Laboratory for Land Forces



UNCLASSIFIED

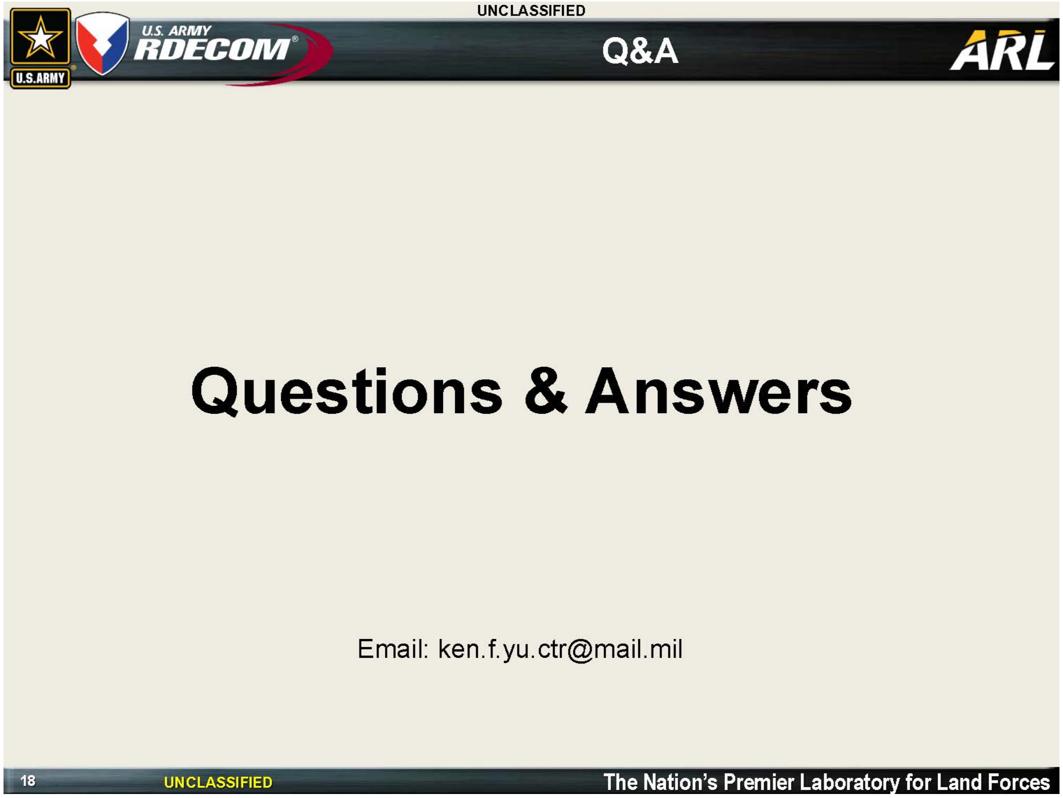
U.S. ARMY RDECOM

Conclusions & Path Forward

ARL

- **Concluding remarks**
  - FAST-D outperforms Snort by examining recent malware variants using outdated Snort signatures
    - FAST-D detected 891 of 905 malwares
    - SNORT detected 545 of 905 malwares
- **Path forward**
  - Examine methods to further improve FAST-D's classification performance WRT both accuracy and resource requirements
    - Enhance multi-class classification for various malware types
    - Reduce FAST-D error rates (e.g., FPR, FNR)
  - Apply FAST-D to other datasets, such as packet header data, adversary stylometry, attack data (e.g., SQL injections), etc.

17 UNCLASSIFIED The Nation's Premier Laboratory for Land Forces



Q&A

Questions & Answers

Email: ken.f.yu.ctr@mail.mil

List of Symbols, Abbreviations, and Acronyms

A	adversarial actions
ACD	active cyber defense
AIS	Adaptive Importance Sampling
APT	advanced persistent threat
ARL	US Army Research Laboratory
ATP	advanced threat protection
AVT	Applied Vehicle Technology
B	adversarial beliefs
COA	course of action
cPPP	Contractual Public Private Partnership
DAGA	Dynamic Adversarial Gaming Algorithm
DARPA	Defense Advanced Research Projects Agency
DBM	deep Boltzmann machine
DNS	Domain Name System
DSB	US Defense Science Board
EBO	Effects Based Operations
EDA	European Defence Agency
ET	Exploratory Teams
FAANG	Facebook, Apple, Amazon, Netflix, and Google
FAST-D	Fast Alert Signature-based Training and Detection
G	adversarial goals
GAFA	Google, Apple, Facebook and Amazon
GIS	geographical information systems
GPR	Generalized plan recognition
HADR	High Availability Disaster Recovery
HFM	Human Factors & Medicine

HTTPS	Hypertext Transfer Protocol Secure
IDS	intrusion detection system
ISIS	Islamic State of Iraq and Syria
IST	Information Systems Technology
LUPI	Learning Using Privileged Information
MANET	Malware and Intrusion Detection for Mobile Ad Hoc Network
MIS	Multiple Importance Sampling
ML	machine learning
MSG	Modelling and Simulation Group
NAT	network address translation
NATO	North Atlantic Treaty Organization
NLP	natural language processing
OFM	operator function model
OODA	observe–orient–decide–act
PA	predictive analytics/analysis
PGG	plan-goal-graph
POV	point of view
RSY	Research Symposium
RTG	Research Task Group
RUNet	closed and controlled internet border
SAR	synthetic aperture radar
SAS	Systems Analysis & Studies
SCI	Systems Concepts & Integration
SET	Sensors & Electronics Technologies
SRA	Strategic Research Agenda
SRIA	Strategic Research and Innovation Agenda
SSM	state space models

STIX	Structured Threat Information Expression
STO	Science & Technology Organization
SURF	Social Understanding and Reasoning Framework
TCP	Transmission Control Protocol
TTP	tactics, techniques, and procedures
UAV	unmanned aerial vehicle
UDP	User Datagram Protocol
X	adversarial axioms

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 CCDC ARL
(PDF) IMAL HRA
RECORDS MGMT
FDCC DCL
TECH LIB

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

5 CCDC ARL
(PDF) FDCC RLC HS
D SHIRES
FDCC RLC HC
E CHIN
FDCC RLW MG
J LENHART
FDCC RLW MA
J SANDS
FCDD RLC ND
T BRAUN