Matching Navy Recruiting Needs with Social Network Profiles Using Lexical Link Analysis

Primary Researcher
Dr. Ying Zhao
408.252.7746
408.218.8484

Associate Researchers:
Dr. Douglas J. MacKinnon
Dr. Shelley P. Gallup

Research Associate Professors
Distributed Information Systems Experimentation (DISE)

Naval Postgraduate School
Matching Navy Recruiting Needs with Social Network Profiles Using Lexical Link Analysis

Naval Postgraduate School, Distributed Information Systems Experimentation (DISE), Monterey, CA, 93943

Approved for public release; distribution unlimited


REPORT TYPE
3. DATES COVERED
4. TITLE AND SUBTITLE
JAN 2010
00-00-2010 to 00-00-2010

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

6. AUTHOR(S)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:
   a. REPORT
   Unclassified
   b. ABSTRACT
   Unclassified
   c. THIS PAGE
   Unclassified

17. LIMITATION OF ABSTRACT
   Same as Report (SAR)

18. NUMBER OF PAGES
   25

19a. NAME OF RESPONSIBLE PERSON

Standard Form 298 (Rev. 8-98)
Proscribed by ANSI Std Z39-18
Research Issues/Scope
Matching Navy Requirements in a Social Networking Environment

Navy Recruiting Requirements

Social Networks
- MySpace
- Facebook
- YouTube
- Twitter

Research Questions

- Do current service entry profiles match well with candidate profiles in the social networking environments?
- What is the methodologies facilitate matching profiles that could potential inform Navy Recruiting Marketing decisions?
- Are there sufficient matches to make marketing in this environment cost effective?
Proposed Method

• **Lexical Link Analysis (LLA)**
  – Features – terms/keywords to describe Navy recruiting needs and candidate profiles
  – Link the features in context of dynamic social network environments, learn from on-going market places

• **Previous applications of (LLA)**
  – System self-awareness/ program-awareness: DISE demonstrated methodology within the Navy acquisition community
    • Matched Navy needs (e.g. need statements) with technologies (e.g. Trident Warrior technologies)
    • Resulting in improved acquisition decisions
    • (See next slide for example)
Lexical Link Maps
(From a previous application)

Specifically: Matched Navy Needs Profiles to Trident Warrior (TW10) Technologies

Overlapping terms or Lexical Links are inside the black box

Distributed Information Systems Experimentation (DISE)  Naval Postgraduate School
Navy Recruiting Research Tasks

**Select and Pre-Process data sets**
- Obtain Facebook data samples - 1,700 Facebook profiles site
- Obtain service entry requirement descriptions and extract features for service entry requirements
- Pre-process and extract features from a sample data set

**Apply Proposed Approaches**
- Profile the current service entry requirements using Lexical Analysis
- Profile Facebook sample data using Lexical Analysis
- Link profiles with rating descriptions using LLA generating potential recruiting candidates
- Develop mapping interface output matching qualified candidates

**Prepare Deliverables**
- NPS Technical Report submitted at the completion of the project
- Presentation at Annual Navy Workforce Research & Analysis Conference in DC, May 2010
- A Trade Journal or Conference Paper submission in the area of Knowledge Management
“Tastes, Ties, and Time (T3): A new social network dataset using Facebook.com”

- Authors
  - Kevin Lewisa, Jason Kaufmana, Marco Gonzaleza, Andreas Wimmerb, Nicholas Christakis

- Affiliations
  - Department of Sociology, Harvard University, United States
  - Department of Sociology, University of California, Los Angeles, United States

- Note: T3 dataset ensures privacy of students’ profiles
Published Research Data

• 1,700 Facebook profiles
  – Data set includes
    • User profiles
    • Interests
    • Communities
    • Forums, group affiliations (schools, colleges and universities, companies, etc.)
    • Friends list, fandom (fan of)
    • Endorsements (supporter of)

• Navy Enlisted Rating descriptions
  – Hard Cards
  – Officer Community Descriptions
Potential Research Benefits

• Improved strategic decisions
• Lexical links can be used to improve
  – Recruiting campaigns
  – Advertisements
  – Regionalization understanding
  – Marketing event strategies
• Real-time marketing
  – Facebook page that takes the feed and disseminates recruiting marketing messages
Potential Impact
Direct Advertising Using LLA Results

Ads can be directed to those who match required characteristics.
Create a dedicated application

*Post marketing information disseminated to fans*
*Grow the fan space*
Ads Placed via Facebook to Matched Profiles & Response Results Tracked
Facebook Profile Data Example

Connect

Interests (36):
adventure, archipelago, cavalry, cold war, dragons, fantasy, fiction, geography, heraldry, history, immortality, longevity, love story, magic, military history, mindpowers, mounted combat, nanotechnology, nanowrimo, national novel writing month, naval history, novel writing, parapsy.

External Services:
  - tan eros@livejournal.com
  - rizanab@...@hotmail.com
  - rizanab@...@gotrillian.com
  - pm_rizanab
  - banazir
  - banazir

Schools:
- Scott Lake Elementary School - Lakeland, FL (1984 - 1985)
- Severna Park Middle School - Severna Park, MD (1985 - 1987)
- Severna School - Severna Park, MD (1987 - 1988)
- Johns Hopkins University - Baltimore, MD (1989 - 1993)
- University of Illinois at Urbana-Champaign - Urbana, IL (1992 - 1998)

Friends [View Entries]

Friends (66):
amishurkier, andrewwyldgonzo, animatortb, banazir, bdm7935, burkhardt, carida_46, cavlee, cheyinka, cloverzi, cretaceousrick, darana, deire, dragnflye, dsthene, elvenwanderer, fyygulicious, gngr, gondhir, gothic_spire, illusio, istani_alal, jadziadax, kaladhwon, klawz, koinonia, idymissas, masaga, maseralida, megruder, majesbar, mirabein, myesko, mom_counsel, mrowe, myng_rabbyt, narri, neobot, nobuddy69, pathothjus, queenmidalal, ringlass, rivendellrose, sahtyinepu, smegalicious, sperose, sui_degneris, tabula_xrasa, taiji_jian, tamf, teremala, the_wanilorn, twinothugin, yodge, zengeneral, zurich31

Data:

- wfdataset.xls: Restricted Access, wave 1 general attributes
- wftbl.net: Restricted Access, wave 1 facebook friends
- wftblgrp.net: Restricted Access, wave 1 housing groupmates
- wftpif.net: Restricted Access, wave 1 picture friends
- wftroom.net: Restricted Access, wave 1 roommates
- wttest.xls: Restricted Access, wave 1 cultural tastes

Distributed Information Systems Experimentation (DISE) 
Naval Postgraduate School
Notional Lexical Link Map
Matching Profiles to Navy Ratings

Social Networker Profiles

Navy Recruiting Needs

Naval History
Adventure
Sailing Ships

Lexical Links

Distributed Information Systems Experimentation (DISE)
Naval Postgraduate School
LLA Recruiting Validation Methods

Figure 1: Validation Model
• **Step 1**: Consider historical sample profiles of recruited personnel, mixed with general populations, for example, samples of youth population in US. The profiles should be similar to the self-reported education and personality characteristics from Facebook.

• **Step 2**: Divide the samples into a validation and a learning set. Set aside the validation set. Use the learning set to match the recruit ratings with the proposed LLA and generate a LLA Match Model.

• **Step 3**: Apply the LLA Match Model to the validation set to recommend the top candidates. Compute the percentage of overlap between the recommended recruits with those who were successfully selected historically (by experts). (Please see the process illustration in Figure 1.) If the overlap percentage or correlation is high, we will consider the matching method as valid.
  – This method can then be used dynamically in automation and thus reduce required manpower.
Rating Descriptions vs. Population Segment

• Data Used
  – Enlisted ratings (AB.doc, AC.doc, AD.doc….)
    • Descriptions
  – JAMRS_2008_SegmentationGuide.pdf
    • 66 segments for 16-20 years of age

• For Validation
  – FY09 National Segmentation Profile.xls
    • Accession rates for all services, for all segments were used for correlation
Result: Lexical Link Table

- **Rows**: Segment description ids
- **Columns**: Enlisted rating ids
- **Number in the cells**: the number of matched word/term categories for a segment and a rating.
Number in the cells: the number of matched word/term categories for a segment and a rating.
# of word categories linked: unique number of categories, i.e. 2, across all the ratings for Segment31
<table>
<thead>
<tr>
<th># of Word Categories Linked</th>
<th>PROGRAMMING--READ OPERATING</th>
<th>SERVE--PEOPLE BATTLEFIELD</th>
<th>SERVE--PEOPLE BATTLEFIELD</th>
<th>SERVE--PEOPLE BATTLEFIELD</th>
<th>PART--WORK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INTERESTED--MECHANICS,PURSUITING</td>
<td>SELF--DISCIPLINE,PACED</td>
<td>SELF--DISCIPLINE,PACED</td>
<td>DEVELOPMENT--SKILL CAREER</td>
<td>PROGRAMMING--READ OPERATING</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>INTERESTED--MECHANICS,PURSUITING</td>
<td>DEVELOPMENT--SKILL CAREER</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>INTERESTED--MECHANICS,PURSUITING</td>
<td>SELF--DISCIPLINE,PACED</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>INTERESTED--MECHANICS,PURSUITING</td>
<td>TECHNOLOGY</td>
</tr>
</tbody>
</table>

**Correlation**

- Pearson correlation between "# of Word Categories Linked" and the accession rate for "ALL SVCS Accessions" from FY09 National Segmentation Profile.xls.
- The correlation of Row 2 vs. 3 is 0.43 (on a scale of 0 to 1)
- Provides Initial validation for this methodology

\[
\rho = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}
\]
Conclusions

- Serves as initial validation LLA method
- We expect that more detailed, individual profiles (e.g. facebook) will increase this correlation
- Improve the accession rate in real-life
- Word/term matches can be used to fit different segment youth to different ratings
References
Collaborative Learning Agents (CLA)
(http://www.quantumii.com)
Internet References