This final report describes the work performed on the GLAD-PC project of the ADAMS program during the final phase of the program (August 1, 2014, July 31, 2015).

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrived as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
This final report describes the work performed on the GLAD-PC project of the ADAMS program during the final phase of the program (August 1, 2014, July 31, 2015).

15. SUBJECT TERMS
machine learning, data analytics, anomaly detection, insider threat detection, quitting detection
ABSTRACT
This final report describes the work performed on the GLAD-PC project of the ADAMS program during the final phase of the program (August 1, 2014, July 31, 2015).

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received Paper

08/19/2013 4.00 Jianqiang Shen, Oliver Brdiczka, Yiye Ruan. A comparison study of user behavior on Facebook and Gmail, ArXiv: 1305.6082, (11 2013): 0. doi: 10.1016/j.chb.2013.06.043

TOTAL: 1

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations
Both of these presentations (see attachments) were given at the final ADAMS PI Meeting in Arlington VA, 4-5 MAR 2015:
(1) Graph Learning and Anomaly Detection using Psychological Context (GLAD-PC)
(2) Technology transition from PARC
Number of Presentations: 2.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received  Paper

TOTAL:
Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>07/14/2015</td>
<td>15.00 Evgeniy Bart, Bob Price, John Hanley. Temporally Coherent Role-Topic Models (TCRTM): deinterlacing overlapping activity patterns, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Database. 07-SEP-15, . : ,</td>
</tr>
<tr>
<td>07/14/2015</td>
<td>14.00 Kumar Sricharan, Gaurang Gavai, Dave Gunning, Rob Rolleston, Mudita Singhal, John Hanley, Juan Julia Liu, Oliver Brdiczka. Detecting employee churn from enterprise social and online activity data, 2015 ASE Eighth International Conference on Social Computing. 18-AUG-15, . : ,</td>
</tr>
<tr>
<td>07/14/2015</td>
<td>16.00 Kumar Sricharan, Gaurang Gavai, Dave Gunning, Rob Rolleston, Mudita Singhal, John Hanley, Juan Julia Liu, Oliver Brdiczka. Detecting insider threat from enterprise social and online activity data, The 7th ACM CCS International Workshop on Managing Insider Security Threats. 12-OCT-15, . : ,</td>
</tr>
<tr>
<td>07/21/2014</td>
<td>10.00 Akshay Patil, Juan Liu, Jianqiang Shen, Oliver Brdiczka, Jie Gao, John Hanley. Modeling Attrition in Organizations from Email Communication, 2013 International Conference on Social Computing SOCIALCOM. 08-SEP-13, . : ,</td>
</tr>
<tr>
<td>08/19/2013</td>
<td>5.00 Jianqiang Shen, Oliver Brdiczka, Juan Liu. Understanding Email Writers: Personality Prediction from Email Messages, UMAP 2013. 10-JUN-13, . : ,</td>
</tr>
<tr>
<td>08/19/2013</td>
<td>9.00 Akshay Patil, Juan Liu, Jie Gao. Predicting Group Stability in Online Social Networks, WWW 2013. 13-MAY-13, . : ,</td>
</tr>
<tr>
<td>08/19/2013</td>
<td>8.00 Hoda Eldardiry, Evgeniy Bart, Juan Liu, John Hanley, Bob Price, Oliver Brdiczka. Multi-Domain Information Fusion for Insider Threat Detection, Workshop on Research for Insider Threat (WRIT). 24-MAY-13, . : ,</td>
</tr>
<tr>
<td>08/19/2013</td>
<td>7.00 Elise T. Axelrad, Paul J. Sticha , Oliver Brdiczka, Jianqiang Shen. Bayesian network model for predicting insider threats, Workshop on Research for Insider Threat (WRIT). 24-MAY-13, . : ,</td>
</tr>
<tr>
<td>08/19/2013</td>
<td>6.00 Francis T. O'Donovan, Connie Fournelle, Steve Gaffigan, Oliver Brdiczka, Jianqiang Shen, Juan Liu, Kendra E. Moore. Characterizing user behavior and information propagation on a social multimedia network, International IEEE Workshop on Social Multimedia Research (SMMR). 15-JUL-13, . : ,</td>
</tr>
<tr>
<td>08/20/2012</td>
<td>2.00 Akshay Patil, Juan Liu, Bob Price, Hossam Sharara, Oliver Brdiczka. Modeling Destructive Group Dynamics in On-line Gaming Communities, International AAAI Conference on Weblogs and Social Media (ICWSM-12). 04-JUN-12, . : ,</td>
</tr>
<tr>
<td>08/20/2012</td>
<td>3.00 Jianqiang Shen, Oliver Brdiczka, Nicolas Ducheneaut, Nicholas Yee, Bo Begole. Inferring Personality of Online Gamers by Fusing Multiple-View Predictions, Conference on User Modeling, Adaptation and Personalization (UMAP 2012). 16-JUL-12, . : ,</td>
</tr>
<tr>
<td>08/20/2012</td>
<td>1.00 Oliver Brdiczka, Juan Liu, Bob Price, Jianqiang Shen, Akshay Patil, Richard Chow, Eugene Bart, Nicolas Ducheneaut. Proactive insider threat detection through graph learning and psychological context, IEEE Workshop on Research for Insider Threat (WRIT). 25-MAY-12, . : ,</td>
</tr>
</tbody>
</table>

TOTAL: 12
(d) Manuscripts

Received Paper


07/21/2014 12.00 Gaurang Gavai, Sricharan Kumar, Juan Liu, Oliver Brdiczka, John Hanley. Predicting Quitting in the Online Yammer Space, The 8th SNA-KDD workshop (06 2014)

TOTAL: 2

Number of Manuscripts:

Books

Received Book

TOTAL:

Book Chapter


TOTAL: 1
Patents Submitted
No new patents were submitted in this last year.

Patents Awarded

Awards
None

Graduate Students

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yiran Wang</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>FTE Equivalent:</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Total Number:</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Names of Post Doctorates

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTE Equivalent:</td>
<td></td>
</tr>
<tr>
<td>Total Number:</td>
<td></td>
</tr>
</tbody>
</table>

Names of Faculty Supported

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTE Equivalent:</td>
<td></td>
</tr>
<tr>
<td>Total Number:</td>
<td></td>
</tr>
</tbody>
</table>

Names of Under Graduate students supported

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTE Equivalent:</td>
<td></td>
</tr>
<tr>
<td>Total Number:</td>
<td></td>
</tr>
</tbody>
</table>
**Student Metrics**

This section only applies to graduating undergraduates supported by this agreement in this reporting period.

- The number of undergraduates funded by this agreement who graduated during this period: 0.00
- The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields: 0.00
- The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields: 0.00
- Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale): 0.00
- Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering: 0.00
- The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense: 0.00
- The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 0.00

---

**Names of Personnel receiving masters degrees**

<table>
<thead>
<tr>
<th>NAME</th>
<th>Total Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Names of personnel receiving PHDs**

<table>
<thead>
<tr>
<th>NAME</th>
<th>Total Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Names of other research staff**

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sri Kumar</td>
<td>0.75</td>
</tr>
<tr>
<td>Gaurang Gavi</td>
<td>0.75</td>
</tr>
<tr>
<td>David Gunning</td>
<td>0.20</td>
</tr>
<tr>
<td>John Hanley</td>
<td>0.25</td>
</tr>
<tr>
<td>Mudita Singhal</td>
<td>0.10</td>
</tr>
<tr>
<td>Eugene Bart</td>
<td>0.10</td>
</tr>
<tr>
<td>Bob Price</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>FTE Equivalent:</strong></td>
<td><strong>2.25</strong></td>
</tr>
<tr>
<td><strong>Total Number:</strong></td>
<td><strong>7</strong></td>
</tr>
</tbody>
</table>
## Sub Contractors (DD882)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 a. Boston Fusion</td>
<td>1 b. 1 Van de Graaff Dr. Ste 107 Burlington MA 01803</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Sub Contractor Numbers (c):**

**Patent Clause Number (d-1):**

**Patent Date (d-2):**

**Work Description (e):** Boston Fusion will provide data analysis support.

**Sub Contract Award Date (f-1):** 7/1/13 12:00AM

**Sub Contract Est Completion Date(f-2):** 6/4/15 12:00AM

## Inventions (DD882)
(1) Foreword
This report details the scientific progress of PARC team in the GLAD-PC project for the period of 08/01/2013-07/31/2014. This research has been funded by the DARPA/ADAMS program under contract W911NF-11-C-0216. Any opinions, findings, and conclusions or recommendations in this report and associated material are those of the authors and do not necessarily reflect the views of the government funding agency.

(2) Table of Contents (None)

(3) List of Appendixes:
• Detecting insider threat from enterprise social and online activity data
• Temporally Coherent Role-Topic Models (TCRTM): deinterlacing overlapping activity patterns
• Detecting employee churn from enterprise social and online activity data
• PARC-ADAMS-PI-Meeting-20150305_v4

(4) Statement of the problems studied
The PARC team investigated three approaches to detecting aspects of malicious insider activity: a) psychological profiling from email; b) quitting dynamics and quitting prediction from corporate social media data; and c) detecting unusual and anomalous behavior from on-line activities.

(5) Summary of the most important results
With regard to (a) Psychological profiling from email: we have defined a Bayesian model for the motivations and psychology of the malicious insider and an associated degree of interest. We aimed then to predict the derived psychological variables automatically from text in emails. Several large studies have been conducted involving over 1000 subjects. We measured the subjects’ psychology using surveys and collected anonymized features from their email communications. We were able to predict the subjects’ psychological variables with up to 95% accuracy (see [Shen1]). The constructed predictors have been applied to various real-world data sets including large corporate email data sets. The results have been made accessible to analysts via a specific personality prediction visualization called the Interactive Personality Workbench (described in last years AUG 2013 – JUL 2014 Interim Report). Initial feedback we received from the analysts is very positive.

With regard to (b) quitting dynamics and quitting prediction from corporate social media data. Last year, we have looked into predicting if and when people quit a corporation using their activity on an internal social media network called Yammer. We got access to a data set of over 24,000 corporate users of this internal social media network of a large corporation, including over 2,000 groups and over 150,000 public messages. The goal was to predict, at any given time instance, if an employee is likely to quit the company. For quitting the company, we have identified 298 quitter instances among 7000 non-quitter instances (after cleaning and filtering the data set according to appropriate parameters, e.g. number of messages and activity scores). Using a random forest and a balanced data sets (50% baseline), we get an accuracy of 68%, which means an improvement of 36% over the baseline. A detailed summary of the results including figures and tables can be in [Gavai1].

During this last year we extended this work quitting dynamics by studying employee churn behavior. Employee churn is a significant concern for organizations, with downsides including loss of talent, its productivity, and also security risk, given that employees are likely to retain confidential company data after they quit. PARC developed hypothesizes that precursors to an employee quitting a company will manifest in the enterprise social and online activity data of the employee. To this end, we processed and extracted relevant features from social data including email communication patterns and content, and online activity data such as web browsing patterns, email frequency, and file and machine access patterns, and used these features to build a predictive model for detecting employee quitting events ahead of time. We tested our predictive models on two different real world data sets, and our experiments show that we are able to detect quitting events with moderately high accuracy. Finally, we build a visualization dashboard that enables managers and HR personnel to quickly identify employees with high quitting scores, which will enable them to take suitable preventive measures to reduce, churn [Sricharan2, attached].

Regarding (c) detecting unusual and anomalous behavior from on-line activities, PARC investigated techniques to discover insider threat in organizations by identifying abnormal behavior in enterprise social and online activity data of employees. To this end, we processed and extracted relevant features that were possibly indicative of insider threat behavior. This includes features extracted from social data including email communication patterns and content, and online activity data such as web browsing patterns, email frequency, and file and machine access patterns. Subsequently, we detect statistically abnormal behavior with respect to these features using state-of-the-art anomaly detection methods, and declare this abnormal behavior as a proxy for insider threat activity. We tested our approach on a real world data set (the Vegas data set from ADAMS) with artificially injected insider threat events. Our experiments show that our proposed approach is fairly successful in identifying insider threat events. Finally, we build a visualization dashboard that enables managers and HR personnel to quickly identify employees with high threat risk scores, which will enable them to take suitable preventive measures and limit security risk [Sricharan1, attached].
PARC also investigated the specific problem of identifying overlapping activity patterns in the VEGAS data set. The Temporally Coherent Role-Topic Model (TCRTM) is a probabilistic graphical model for analyzing overlapping, loosely temporally structured activities in heterogeneous populations. Such loose temporal structure appears in many domains, but especially in the ADAMS data, where individual events that make up an activity have coherence, but no strong temporal ordering. For instance, preparing a PowerPoint presentation may involve opening files, typing text, downloading images, and saving files. These activities occur together in time, but without a strong ordering or fixed duration. These temporally coherent activities may also overlap—the user might also be responding to email while working on the presentation. Finally, the population of users has subgroups—in the office, administrators, salespeople and engineers will have different activity distributions. The unique architecture of the TCRTM model allows it to automatically infer an appropriate set of roles and activity types while simultaneously assigning users to these roles and segmenting their event streams into high-level activity instance descriptions. On two real-world datasets taken from computer user monitoring and social services debit card transactions we show that TCRTM extracts semantically meaningful structure and improves perplexity score on hold-out data by a factor of five compared to standard models such as LDA [Bart1, attached].

All of these results and summary of PARC’s work on ADAMS was presented at the final ADAMS PI Meeting, held at DARPA, in March 2015 (the briefing slides are attached).

(6) Bibliography
Books

Peer-Reviewed Conference Proceeding publications (other than abstracts)
Kumar Sricharan, Gaurang Gavai, Dave Gunning, Rob Rolleston, Mudita Singhal, and John Hanley, Detecting insider threat from enterprise social and online activity data, the 7th ACM CCS International Workshop on Managing Insider Security Threats, 12-16 OCT 2015.


Kumar Sricharan, Gaurang Gavai, Dave Gunning, Rob Rolleston, Mudita Singhal, John Hanley, Juan Julia Liu, and Oliver Brdiczka, Detecting employee churn from enterprise social and online activity data, 2015 ASE Eighth International Conference on Social Computing, 18-20 AUG 2015.


Jianqiang Shen, Oliver Brdiczka, Juan Liu. Understanding Email Writers: Personality Prediction from Email Messages, UMAP 2013. 10-JUN-13, .

Jianqiang Shen, Oliver Brdiczka, Nicolas Ducheneaut, Nicholas Yee, Bo Begole. Inferring Personality of Online Gamers by Fusing Multiple-View Predictions, Conference on User Modeling, Adaptation and Personalization (UMAP 2012). 16-JUL-12, .


Manuscripts submitted, but not published
Gaurang Gavai, Sricharan Kumar, Juan Liu, Oliver Brdiczka, John Hanley. Predicting Quitting in the Online Yammer Space, The 8th SNA-KDD workshop (06 2014)
Papers published in peer-reviewed journals
Jianqiang Shen, Oliver Brdiczka, Yiye Ruan. A comparison study of user behavior on Facebook and Gmail, ArXiv: 1305.6082, (11 2013): 0. doi: 10.1016/j.chb.2013.06.043
Technology Transfer

See attachment: Technology transition from PARC
“Graph Learning and Anomaly Detection using Psychological Context (GLAD-PC)

The PARC team, 03/05/2015
### The Team

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Skills/Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>David Gunning</td>
<td>PI</td>
<td></td>
</tr>
<tr>
<td>John Hanley</td>
<td>Info. Architecture</td>
<td>Database</td>
</tr>
<tr>
<td>Mudita Singhal</td>
<td>Visualization Design</td>
<td></td>
</tr>
<tr>
<td>Oliver Brdiczka</td>
<td>Quitting Example</td>
<td></td>
</tr>
<tr>
<td>Sricharan Kumar</td>
<td>Machine learning</td>
<td></td>
</tr>
<tr>
<td>Gaurang Gavai</td>
<td>Info. Architecture</td>
<td>Feature Extraction</td>
</tr>
<tr>
<td>Rob Rolleston</td>
<td>Visualization Design</td>
<td>&amp; Implementation</td>
</tr>
<tr>
<td>Julia Liu</td>
<td>Quitting Example</td>
<td></td>
</tr>
</tbody>
</table>

### Subcontractors

**Boston Fusion**
Remote operations, group dynamics analysis

- Betsey Benagh
- Joanna Brown
- Connie Fournelle
- Kendra Moore
Personality and Anomalous Behavior

- Organizations (and society) face increasing amount of threats from “inside” and “outside”
- Challenge: Uncover malicious behavior in a *timely* way through automatic analysis
- Anomalous behavior trace often precedes the actual “incident”
- Personality has been shown to be a reliable indicator for future (malicious) behavior (Jaclyn et al., 2011)
- 50% of job quitters steal confidential company data
• PARC project:
  - Graph Learning and Anomaly Detection using Psychological Context (GLAD-PC)
  - Idea: combine graph learning / structural anomaly detection and psychological modeling
Previous Research: Personality Profiling for Malicious Insider Detection

• We are interested in **psychological profiles** as indicators for future malicious behavior

• Why?
  
  • Counterproductive (cyber-)behaviors have been shown to be **highly correlated** with Big-5 personality variables [1]
  
  • Actual insider threats have a low base rate → psychological profiles are a powerful filter to reduce false positives

---

What is a Personality Profile?

Personality Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>20 to 35</th>
<th>35 to 45</th>
<th>45 to 55</th>
<th>55 to 65</th>
<th>65 to 80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>4.56</td>
<td>24.1</td>
<td>42.6</td>
<td>24.1</td>
<td>4.56</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>7.64</td>
<td>23.5</td>
<td>36.1</td>
<td>24.6</td>
<td>8.15</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>50.2 ± 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excitement Seeking</td>
<td>50 ± 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostility</td>
<td>4.56</td>
<td>22.0</td>
<td>40.7</td>
<td>26.2</td>
<td>6.45</td>
</tr>
<tr>
<td>Extraversion</td>
<td>50.8 ± 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Assurance</td>
<td>50 ± 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Mood/Emotion</td>
<td>4.95</td>
<td>24.1</td>
<td>41.7</td>
<td>24.2</td>
<td>4.98</td>
</tr>
<tr>
<td>Organizational Deviance</td>
<td>4.79</td>
<td>24.0</td>
<td>42.1</td>
<td>24.3</td>
<td>4.78</td>
</tr>
<tr>
<td>Personal Deviance</td>
<td>7.31</td>
<td>51 ± 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>23.3</td>
<td>36.8</td>
<td>24.7</td>
<td>7.96</td>
<td>50.3 ± 12</td>
</tr>
</tbody>
</table>
Approach

- **Idea:** automatically estimate personality from emails
Results

• **Data Collection & Evaluation Results (of Estimators):**
  – Over 1000 personality profiles + emails collected from MechTurk and company-internal (for training the estimations)
Interactive Personality Visualization

User: user0010
x-Axis: Hostility ▼ y-Axis: Perceived Stress ▼

<table>
<thead>
<tr>
<th>Agreeableness</th>
<th>Conscientiousness</th>
<th>Extraversion</th>
<th>Hostility</th>
<th>Interpersonal Deviance</th>
<th>Neuroticism</th>
<th>Overall Mood</th>
<th>Perceived Stress</th>
<th>Self Assurance</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.5</td>
<td>30</td>
<td>30.5</td>
<td>32.8</td>
<td>49.8</td>
<td>45.1</td>
<td>13.9</td>
<td>28.2</td>
<td>49.8</td>
</tr>
</tbody>
</table>

![Graphs showing data points and histograms for different personality traits and stress levels.](image-url)
Previous Research: Quitting and Destructive Group Dynamics

• We are interested in quitting behavior & destructive group dynamics

• Proxy of malicious behavior: “50% job leavers steal confidential company data”

• Questions:
  – Can we observe quitting behavior and destructive group dynamics in real-world and social space
  – How is real-world behavior related to social space data
  – Can we predict real-world behavior
Previous Research on Quitting Behavior

### Online Games

**Destructive group dynamics**
- if/when a player will quit a guild
- damage associated with a quit event
- guild stability against member loss

### Startup Venture

**Churn prediction in a real-world corporation**
Predict quitting based on work practice, email, and content.

- **Database** ~180,000 emails, up to 2yr observation
- **Work practice features**
- **Email graph features**
- **Content features**
- **Correlation**

### Yammer

**Real-world corporation**
Voluntary Turnover (Quitting) • Recent quitters • N=12 (Male = 9, Female = 3) • Job titles include: • research scientist (2) • software engineer (3) • research engineer (2) • director/manager (1) • senior associate in a bank (1) • system engineer (1) • manufactory engineer (1) • office manager (1)
Interview Results

• Web Browsing
  – Increased use of career sites (e.g., LinkedIn)
  – Increased browsing of company profiles

• Personal Email
  – Increased use of personal email for job applications

• Work Email
  – No conscious change
  – Some made an effort to maintain normal email behavior

• Work Routine
  – Shortened work hours and more time off to accommodate interviews

• Multitasking
  – Shortened attention spans at work
  – more task switching

• Engagement
  – Decreasing engagement in general
  – More neutral sentiment in emails
Features Selection/Engineering

- Extract a rich set of features:
  - Email Usage (-sent count)
  - Email Content (-subject char length)
  - Log On / Log Off Statistics
  - Application Activity (+max time spent on activity, + # of activity types per day)
  - Web Usage (time on –internal/+job sites)
  - Feature matrix F: U x T x D
This Year’s Problem Set-up

Vegas Database

Features
- Email Usage
- Email Content
- Logon/Logoff
- App. Activity
- Web Usage

Quitting Examples

Quitting Classifier

Anomaly Detector

Quitting Prediction

Anomaly Prediction
Problem set-up

- Twin approaches:
  - Supervised – Use quitting labels as proxy
    - Build classifier to predict quitters and corresponding time instances
  - Unsupervised – Use anomaly detection methods to detect abnormal behavior
Vegas Dataset

- Multi-Domain Employee Data
- Anonymized application-wise log of User activity
- Anonymized activity log of user interactions with different agents
- Email interaction data between business unit users
- Aggregated statistics on Email content data
- Snapshots of LDAP hierarchy
- Day-to-day LDAP diffs
## Vegas Dataset

<table>
<thead>
<tr>
<th>Dataset Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Date range</strong></td>
</tr>
<tr>
<td><strong>Users</strong></td>
</tr>
<tr>
<td><strong>Dataset Size</strong></td>
</tr>
<tr>
<td><strong>Domains</strong></td>
</tr>
</tbody>
</table>
| **Target Users**   | - 555 Quitters (1270 Pseudo)  
|                    | - 104 Red Team Users |
Feature Extraction

- Calculated aggregate features from raw data
- Constructed features in 5 different domains
- Features developed from earlier Yammer work were supplemented with newer ones derived from insights gained by conducting interviews with employees that quit their jobs

<table>
<thead>
<tr>
<th>Email Usage Features</th>
<th>Logon Logoff Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly sent count</td>
<td>Number of logons</td>
</tr>
<tr>
<td>Weekly read count</td>
<td>Number of logoffs</td>
</tr>
<tr>
<td>Number of messages sent in the day</td>
<td>Number of hours with logon activity</td>
</tr>
<tr>
<td>Number of messages sent at night</td>
<td>Number of hours with logoff activity</td>
</tr>
<tr>
<td>Number of messages read in the day</td>
<td>Activity Features</td>
</tr>
<tr>
<td>Number of messages read at night</td>
<td>Number of activity types</td>
</tr>
<tr>
<td>Email Content Features</td>
<td>Max contiguous time spent on activity</td>
</tr>
<tr>
<td>Average subject word length</td>
<td>Number of activities</td>
</tr>
<tr>
<td>Average subject character length</td>
<td>Time spent on on email applications</td>
</tr>
<tr>
<td>Average content character length</td>
<td>Time spent on on productivity applications</td>
</tr>
<tr>
<td>Average content word length</td>
<td>Time spent on on web applications</td>
</tr>
<tr>
<td>Average content sent length</td>
<td>Time spent on on engineering applications</td>
</tr>
<tr>
<td>Number of exclamation points</td>
<td>Web Usage Features</td>
</tr>
<tr>
<td>Number of multiple exclamation points</td>
<td>Time spent on on websites</td>
</tr>
<tr>
<td>Number of question marks</td>
<td>Time spent on on career sites</td>
</tr>
<tr>
<td>Number of multiple question marks</td>
<td>Time spent on on web mail sites</td>
</tr>
<tr>
<td>Number of brackets</td>
<td>Time spent on on entertainment sites</td>
</tr>
<tr>
<td>Number of dashes</td>
<td>Time spent on on internal SM sites</td>
</tr>
<tr>
<td>Number of double dashes</td>
<td>Time spent on on internal sites</td>
</tr>
<tr>
<td>Number of ellipses</td>
<td>Time spent on on news sites</td>
</tr>
<tr>
<td>Number of commas</td>
<td>Time spent on on private social media sites</td>
</tr>
<tr>
<td>Number of semicolons</td>
<td>Time spent on on search sites</td>
</tr>
<tr>
<td>Number of colons</td>
<td>Time spent on on tech sites</td>
</tr>
</tbody>
</table>
Hierarchy Creation

- Needed a hierarchy of the organization to be able to compare the behavior of a user with their peers
- Data available: daily snapshots of LDAP hierarchy
- We created a normalized hierarchy by finding the most persistent relationships between supervisors and employees over the time period in consideration
- Resulted in ~200 sub-trees due to the business unit not containing the higher levels of the hierarchy
Hierarchies

- Examples of sub-trees
- ext---- nodes are external to the business unit
Problem statement: At any given time, predict if an employee is likely to quit the company:

- Restrict attention to (User U, Time T) tuples such that user U has data for at least 1 month leading up to time T
- 0.6M such total instances; 2K / 0.6M (~ 0.5%) instances are when user U has quit in time T, T-1 or T-2
- Subsample to deal with class-imbalance problem
Supervised approach
Quitting Detection

• Accuracy = 73% using Random forests (46% improvement compared to random baseline)

• Content features are most predictive for quitters and pseudo-quitters

• Confusion Matrix:

<table>
<thead>
<tr>
<th>Class</th>
<th>+</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>0.746</td>
<td>0.254</td>
</tr>
<tr>
<td>-</td>
<td>0.310</td>
<td>0.690</td>
</tr>
</tbody>
</table>
Supervised approach

Quitting Detection

Receiver operating characteristic for quitting prediction

- ROC curve (area = 0.76)
Quitting Visualization Dashboard

ADAMS Dashboard

Quitting Risk and Anomaly Detection Features Using 4,524 out of 4,524 records | Reset All

Average values over date range: 2013-10-07 to 2014-06-30

<table>
<thead>
<tr>
<th>user6</th>
<th>Quitter</th>
<th>Pseudo</th>
<th>RedTeam</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>06F2A3</td>
<td>true</td>
<td>true</td>
<td>false</td>
<td>0.21</td>
</tr>
<tr>
<td>142F9E</td>
<td>true</td>
<td>true</td>
<td>false</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Quitting Visualization Dashboard
Supervised approach
Quitting Detection - Insight

- Quitting scores tend to peak ~2 weeks before quitting
Supervised approach
Quitting Detection - Insight

- Quitting scores tend to peak ~2 weeks before quitting
Unsupervised approach
Quitting Detection

Detect anomalies with respect to two aspects:
• Detect if user is anomalous with respect to rest of the employees at each time instance
• Detect if user’s behavior has changed drastically over time
• Idea: In addition to features F, also construct differences
  \[ dF = F[:,T+1,:] - F[:,T,:] \]
• Run iForest on joint matrix \([F;dF]\)
Unsupervised approach
Quitting Detection
Unsupervised approach
Quitting Detection

- Can identify 46% of red-team events by tracking top 15% of users every week
- 85% by tracking top 35%
Anomaly Visualization Dashboard

Quitting Risk and Anomaly Detection Features Using 4,524 out of 4,524 records | Reset All

Average values over date range: 2013-10-07 to 2014-06-30

<table>
<thead>
<tr>
<th>user6</th>
<th>Quitter</th>
<th>Pseudo</th>
<th>RedTeam</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>03C576</td>
<td>false</td>
<td>false</td>
<td>true</td>
<td>0.44</td>
</tr>
<tr>
<td>3671A4</td>
<td>false</td>
<td>false</td>
<td>true</td>
<td>0.46</td>
</tr>
<tr>
<td>7E75F8</td>
<td>false</td>
<td>false</td>
<td>true</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Anomaly Visualization Dashboard
Conclusion and Future Work

- End-user activity can be used to determine suspect insider threat behavior

- False-alarms fairly significant, due to
  - Rarity of abnormal events
  - Statistical anomalies that do not translate to real world
Conclusion and Future Work

• Further research needed to bring down false alarm rate
  • Integration of external data sources
  • Integration of psychological modeling
  • Incorporating analyst feedback to select features
Thank you!

Questions?