Insider Threat Detection
A Heuristic Network-Based Approach

Prof. Kathleen M. Carley
CASOS, ISR, SCS, Carnegie Mellon University

Geoffrey Morgan
CASOS, ISR, SCS, Carnegie Mellon University

Nikhil Behari
CASOS Intern

An insider threat is an individual or group with access to an organization's essential network, systems, or data, that deliberately abuses their access for malicious purposes. The 2014 US State of Cybercrime survey states that up to 46% of electronic crimes were committed by insiders, and 37% of cases could not be referred for legal action because the organization could not identify the individual or individuals responsible for the cybercrime. Therefore, the aim of this study was to determine the utility of five different communication-based feature sets through the use of a rule-based machine learning algorithm to detect and mitigate insiders within an organization.

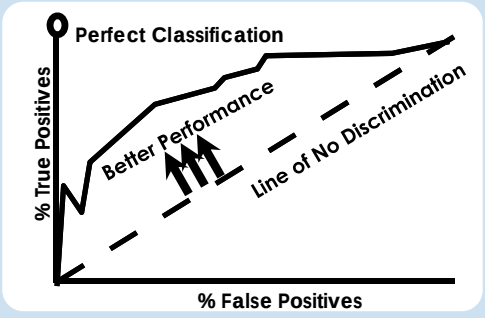
Methodology

1) Obtain Enron Corpus

2) Corpus to Network Transformation

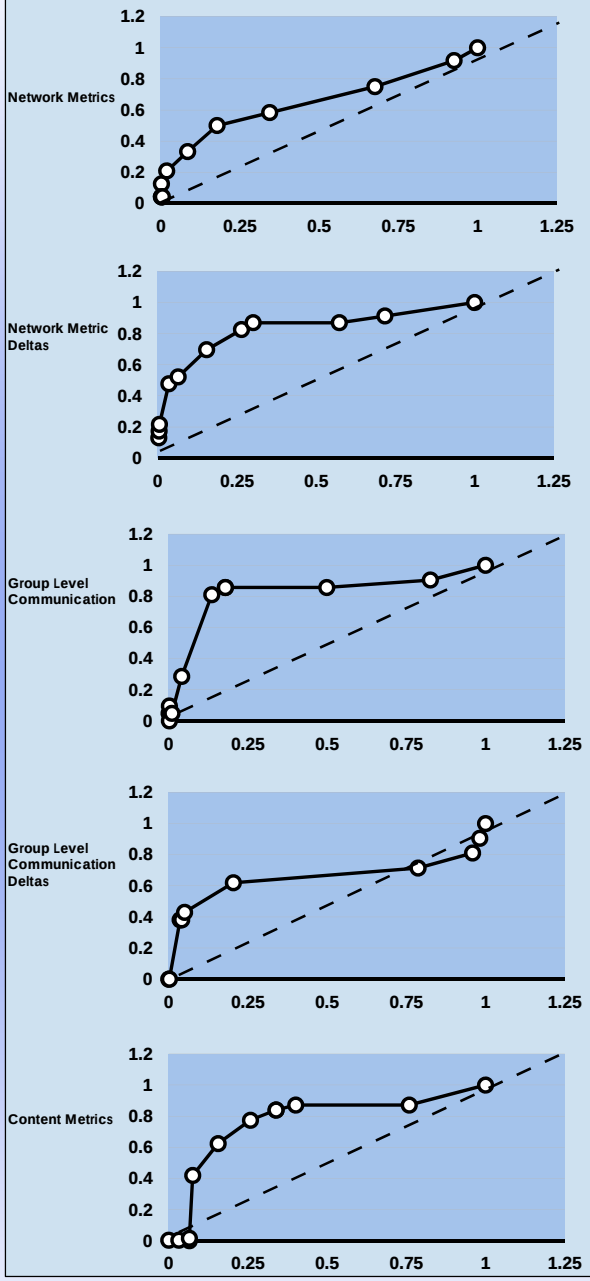
3) Feature Sets Extraction
   - Network Metrics
   - Network Metric Deltas
   - Group-Level Communication Features
   - Group-Level Communication Deltas
   - Content Metrics

4) Machine Learning and Analysis of Results
- JRip in Weka was used to create receiver operator characteristic curve for each feature set through five-fold cross validation
- Cost of missing insider was adjusted through cost-sensitive meta-classifier to produce table of false vs. true positives
- ROC curve produced to visualize utility of each feature set



Sample ROC Curve demonstrating classifier effectiveness

Results



This research presents a novel approach to insider threat detection through social network analysis and supervised machine learning, and demonstrates the utility of several feature sets that may be useful in insider threat mitigation.

Part of this work was done by Nikhil Behari as part of his highschool science fair project. In addition, the effort by Dr. Carley and Geoff Morgan was supported in part by the Office of Naval Research (ONR) through a MURI N000140811186 on adversarial reasoning, CERT, and the Center for Computational Analysis of Social and Organization Systems (CASOS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research, CERT, or the U.S. government. The authors would like to thank Andrew Moore at the Software Engineering Institute/CERT for guidance and feedback.