Phishing Detection Using Semi-Supervised Methods with New Features

Victor Zeng
Advisor: Rakesh M. Verma
Motivation

- Phishing is the act of sending fake emails to trick a user into doing something.
- Beachhead for 95% of attacks on enterprise networks
- Average cost: $1.6 Million
- Cannot depend on user to identify phishing emails
- Creating labeled training data is expensive

Goal

- Improve upon the current state-of-the-art THEMIS model
- Publish a paper based on my results
Objectives

• Identify new features which can be used for phishing detection
• Use semi-supervised methods to detect phishing emails
Expected Impact

• Improve performance of phishing detection methods
• Decrease the amount of labeled data required to train phishing detection models
Deliverables

• Code + Documentation
• Poster
• Report
• Paper
• Final Presentation
Methods: Objective 1

Perform exploratory analysis on proposed feature

Implement feature in PhishBench 2.0

Evaluate multi-feature performance with PhishBench 2.0

Evaluate single-feature performance with PhishBench 2.0
Results: Objective 1

- Spellcheck ratio feature
  - Statistically different between phish and legit emails (p-value: $1.512 \times 10^{-22}$)
  - Random Forest identifies 54% of phish emails in single feature test
Conclusions

• Spellcheck ratio is a promising feature for phishing detection
Methods: Objective 2

1. Extend PhishBench 2.0 to support semi-supervised methods
2. Implement semi-supervised methods in PhishBench 2.0
3. Evaluate performance of semi-supervised methods against pre-existing supervised methods
Remaining Work

- Evaluate features from Statement Analysis
- Acquire additional datasets
- Work for Objective 2
The REU project is sponsored by NSF under award NSF-1659755. Special thanks to the following UH offices for providing financial support to the project: Department of Computer Science; College of Natural Sciences and Mathematics; Dean of Graduate and Professional Studies; VP for Research; and the Provost's Office. The views and conclusions contained in this presentation are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors.