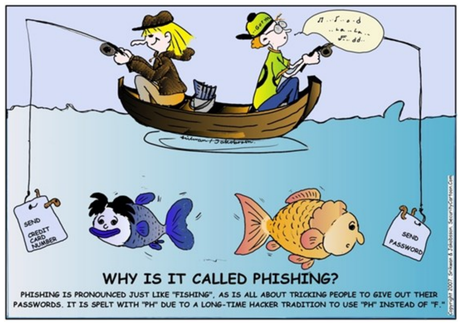
Phishing Detection with Deception Features

ReDAS Lab

Victor Zeng, Dr. Rakesh M. Verma, University of Houston

Problem

Phishing is convincing a victim to perform some action through deception over email. It is normally used to acquire financial information or passwords, but it can also be used to deliver malware. The average cost of a phishing attack for a mid-sized business is $1.6 million dollars, and they form the initial beachhead for 95% of attacks on enterprise networks.



Goal

Identify new features that are useful for identifying phishing emails by machine learning.

Dataset

In this research, we used the IWSPA 2018 dataset.

Method

Perform Basic Statistical Analysis of Feature

↓

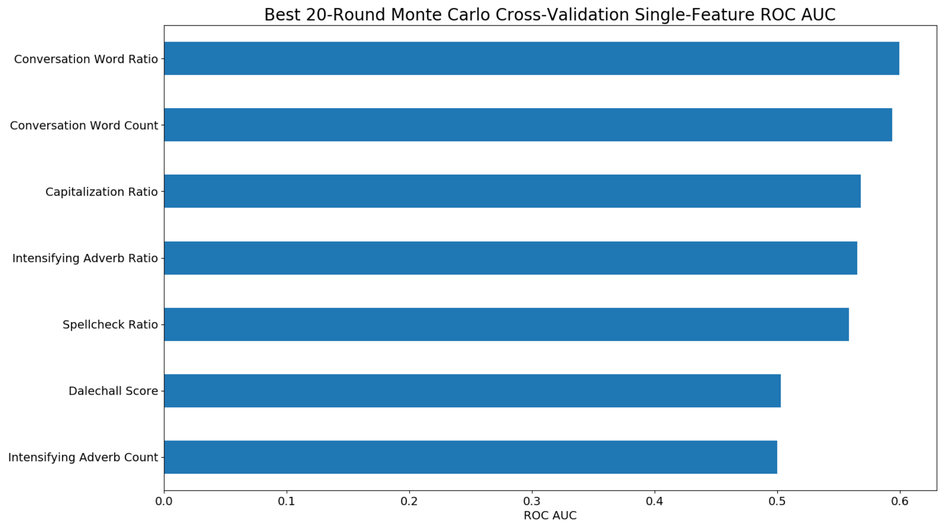
Perform Single Feature Experiment

↓

Rank Features by ROC AUC Score

Single Feature Experiment— 20 Round Monte Carlo Cross Validation using the cleaned IWSPA dataset.

Results



Discussion

- We successfully identified several new features that are useful for phishing detection.
- Counting features performed better when normalized by the length of the email.
- Clarity metrics such as the Dalechall Score perform surprisingly poorly given the t-test results.

Future Work

- We plan to continue developing and testing new features for phishing detection.
- Many of the features evaluated involve counting instances of key words in the text. In the future, we would like to develop a method to automatically identify these key words.

This research was made possible by NSF grants to the University of Houston Computer Science Department (NSF CNS-1551221, NSF CNS 1319212).

COMPUTER SCIENCE