Detection of phishing pages based on interaction

Phishing attacks are growing drastically in the last years. The need to automatically detect phishing sites becomes critical. In this project we wish to implement a number of methods that interactively communicate with site suspected of phishing and decide the probability of it being a phishing site or not, based on how the site behaves.

Phases

1. Build dataset of benign & phishing sites. You can use Phishtank as a source for phishing sites and the actual phished services as the benign data set.
2. Define the interactions we wish to test.
3. Execute the interactions - what are the differences between benign and phishing sites responses. Document the results.

Deliverables

1. All data collected from the research including url’s, request data and response data.
2. A list of possible interactions with the suspected web site with the differences between benign and phishing sites for each interaction.
3. The code developed for testing the interactions.
4. A report which includes a summary of what was done in each step of the project and:
   a. What is the statistical evidence for the decisions you got for the algorithm
   b. What were the algorithm results based on your testing
   c. What is your recommendation going forward

Avishay Bartik
avishayb@il.ibm.com
050-2133732

Aviv Ron
rona@il.ibm.com
058-5810200