Computer Security Homework 5 Report

Kaleb Lott, Jazlyn Maxwell, Connor Prinster, and Ben Taylor

1 December 2019

Phishing

The likelihood of a Phishing attack is represented as a percentage. The percentage is increased by instances of the following suspicious behavior: spear-phishing, inclusion of URLs, titles of authority, threats of time-sensitivity and consequences, and reward offers. Each instance of a suspicious behavior is counted, and the total number of instances of a specific behavior are weighted according to how closely those behaviors are associated with phishing attacks. In our system, we determined suspicious URLs, email addresses, and threats of consequences to be most closely associated with phishing attacks; while spelling errors are loosely associated with phishing attacks. The average of each weighted sum is taken and converted into a percentage. The largest weighted value is reported as the largest threat.

The phishing security scanner could be vastly improved by implementing machine learning algorithms. The current scanner has no ability to take context into account. Machine learning algorithms could be used to perform a sentiment analysis on the content of emails in order to account for context. Algorithms could also more effectively search for spoofed email addresses and URLs.

SQLI

The likelihood of an SQLI attack is represented as a percentage, with a maximum of 100%. The percentage is increased by instances of the following suspicious behavior: comments, alternate encodings, apostrophes, inclusion of risky statements (such as union, from, or select), spaces, tautologies, illegal or incorrect queries, inferences, and piggy-backing. Each instance of a suspicious behavior is counted, and the total number of instances of a specific behavior are weighted according to how closely those behaviors are associated with SQLI attacks. In our system, we determined alternate encodings, comments, risky statements, and piggy-backed statements to be most closely associated with SQLI attacks. The sum of the weighted values is taken and converted into a percentage. The largest weighted value is reported as the largest threat.

Our scanner only attempts to identify SQLI attacks based on the source code. The scanner could be improved by running the SQL code in a controlled instance of the database in order to see what it does. The current approach can never catch all SQLI attacks, while running the SQL code could find more instances of malicious code. Machine learning can be used to determine whether or not possibly benign behaviors, such as comments in a query, are hiding an SQLI attack.