CALIFORNIA STATE SCIENCE FAIR
2012 PROJECT SUMMARY

Name(s)

Eric Sauer

Project Number

S1422

Project Title

PDFClearance: Developing Software to Detect Malicious PDFs

Abstract

Objectives/Goals

PDFs are a globally used file format making them a new chosen vector for malware attacks. Because PDFs can run JavaScript, hackers have learned how to exploit these files and hide their attacks deep within the code. Currently there is no anti-virus software that states clearly whether an unopened PDF, containing a new or altered malicious PDF exploit, is malicious. The purpose of this research project was to develop a new type of software called PDFClearance that detects the behavioral intent of an unopened PDF file by analyzing the JavaScript methods within its text.

Methods/Materials

Over 1100 malicious and benign PDFs were analyzed by comparing their uses of JavaScript and seven malicious uses were determined. PDFClearance looks for these malicious uses of JavaScript and determines whether an unopened PDF file is most likely benign, possibly malicious, or most likely malicious. The new software was tested for its performance with the previously analyzed collection of 1100 PDFs to confirm successful identification of the seven malicious JavaScript uses and a random selection of 1000 PDFs from a not yet analyzed PDF data base to determine its accuracy in correctly identifying malicious PDFs from benign ones.

Results

It was hypothesized that if malicious uses of JavaScript in PDFs can be determined, then it is possible to create software that can detect whether PDFs are malicious with at least 95% accuracy. From the data obtained PDFClearance was successful in identifying a malicious PDF 96% of the time and a benign PDF 97% of the time, supporting the research hypothesis. When identifying PDFs from the random collection of benign and malicious PDFs, PDFClearance correctly identified PDFs 95% of the time.

Conclusions/Discussion

PDFClearance was successful in correctly identifying the behavioral intent of an unopened PDF. This new software offers a significant improvement for cyber defense against malicious PDFs and helps protect the information and systems we rely on every day contributing to the cyber security of our nation. Future development will involve incorporating PDFClearance into current PDF-readers and also creating a web browser plug-in to check PDFs as they are downloaded.

Summary Statement

This project analyzes a collection of malicious and benign PDFs and develops new software that can detect the behavioral intent of an unopened possibly malicious PDF.

Help Received

Mentor Dr. Thomas Kroeger from Sandia National Laboratory answered questions I had about cyber defense issues.