Initial Plan

Identifying the best machine learning model for web-based attack

Cardiff University School of Computer Science and Informatics

Author: Alice Edwards

Supervisor: Amir Javed

Contents

One Semester Individual Project – 40 credits

CM3203

Description

The aim of this project is to develop a machine classifier to detect attacks on web applications. To do this I am building a vulnerable web application and carrying out various attacks on it, with the intent to use machine learning to classify network traffic into either benign or malicious.

A Web application (Web app) is an application program that is stored on a remote server and delivered over the Internet through a browser interface. [1] Any interaction with a web app will go through a server. If a server was compromised from an attack any device that uses it could be affected. Most websites have web applications. When a first responder is called on to examine a suspected attack on a device their first responsibility is to determine whether there has actually been an attack. This machine classifier should help with this as it identifies the web traffic in the context of a web application.

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves. [2]

This application allows for users to monitor their network traffic and be alerted when they are at risk due to the way in which network traffic changes. To do this I will need a large sample of network traffic to highlight where an attack changes the nature of the traffic.

I will be using OWASP Broken Web Applications Project to support me in creating a vulnerable web application. Through my attacks I will intercept traffic between the application and the server, so that I can then use the intercepted network traffic to build a machine classifier to detect cyber-attacks on malware based on network traffic.

Ethics

This project deals with handing network traffic. However, I do not believe any personal data will be stored in the network traffic / accessible to me. Therefore, I do not need ethical approval. However, if anything is to change on this I will immediately file for ethical approval, having already completed the course previously.

I will also throughout the project view all decisions in regard to their ethical impact. I will review the ethics of all decisions made in my final report and ensure to the best of my ability all laws and guidelines are adhered to, especially in regard to data.

Project Aims and Objectives

Background Objectives:

- Establish an understanding of different types of attacks on network traffic
    - Produce a list of possible attack vectors I can use on my web application
- Establish an understanding of different vulnerabilities a web application can have and why it makes them insecure

Primary Aims: The minimum requirements:

- I will develop a working web application using known vulnerabilities
- I will record and document attacks run on the web application
- To develop, through coding a machine learning application, a network classifier
    - The machine classifier must be able to take network traffic as an input
- From the input it should access using its learning on the type of traffic it is – whether it is malicious or not

Secondary Aims: Additional requirements, that would be nice to implement:

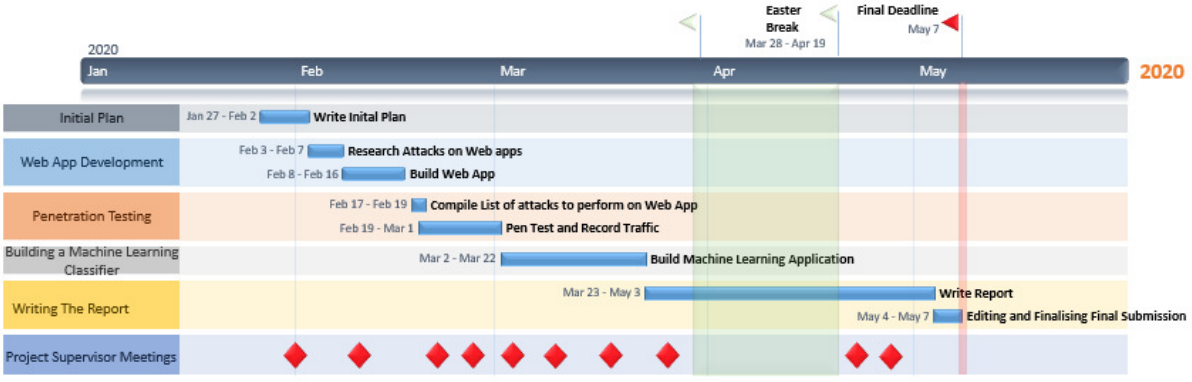
- Where malicious traffic is detected the classifier may be able to identify what kind of attack is taking place beyond it being malicious

Work Plan

Broad plan:

Week	Objective	Milestone
1 27/01 – 02/02	Complete Initial Plan Skype Meeting with Supervisor	Submit Initial Plan by 03/02
2 03/02 – 09/02	Research Attacks Build Web App	Start Web App Development
3 10/02 – 16/02	Skype Meeting with Supervisor	Finish Web App Development by 16/02
4 17/02 – 23/02	Compile list of all attacks to perform on web app & perform pen testing on the web app	Begin Pen Testing
5 24/02 – 01/03	Meeting with Supervisor x 2	Finish Pen Testing by 02/03
6 02/03 – 08/03	Build Machine Learning Application Meeting with Supervisor x 3	Begin Machine Learning
7 09/03 – 15/03		
8 16/03 – 22/03		Finish Machine Learning Application by 22/03
9 23/03 – 29/03	Report Writing Meeting with Supervisor x 3	Begin Writing Report
Easter 28/03 – 19/04		
10 20/04 – 26/04		
11 27/04 – 03/05		Finish Writing Report by 03/05
12 04/05 – 07/05	Proof Reading and Submission of Report	Submit Final Report

Gantt Chart:



Works Cited

[1] Search Software Quality, “Web application (Web app),” August 2019. [Online]. Available: https://searchsoftwarequality.techtarget.com/definition/Web-application-Web-app. [Accessed January 2020].

[2] E. S. Team, “What is Machine Learning? A definition,” 7 March 2017. [Online]. Available: https://expertsystem.com/machine-learning-definition/. [Accessed January 2020].