Sean Sanders

RIT

Mathematica and wireshark Security Analysis

**INTRODUCTION**

Mathematica is a powerful tool that can be used to analyze Wireshark file captures and enable administrators to detect suspicious traffic on the network. This tool allows users to use multiple programming languages and solutions all in a single environment. The utilization of Mathematica enables network administrators to detect possible flaws or vulnerabilities by importing Wireshark capture files into the environment. Another benefit of Mathematica is the integration of map reduce tasks directly into the environment, eliminating clutter and confusion when viewing large amounts of data. This computer program renders the use of Hadoop unnecessary, for the task can be performed locally on your own system.

Project plan.

What I would like to do is create a tool that will be able to analyze a Wireshark file and extract important information to increase the detection of a security threats or suspicious activity on the network. In doing so I will need to use Wireshark to capture the packets that cross the network. The packet capture will consist of an hour or two worth of data from a pod in the networking lab. This data will then be saved to GitHub because then it will be accessible from any device and revert any changes if the file is changed. GitHub will also be the means by which all data will be backed up. The next step will be to import data into Wolfram Mathematica. After importing the data into Mathematica I will have to search for key criteria such as Bit Torrents, Bitcoin, and Netflix in order to narrow the search for security threats to the network.

The reason that Netflix can be considered a security threat is because of the vast amount of data crossing the network, which threatens the quality of service. Bit torrents are also a threat because they can contain malicious malware that can harm the PC. Likewise, bitcoin transactions on a network can be deemed dangerous because they are a digital currency that could be stolen through online packet captures. In general, a security threat can be anything related to suspicious interaction with a network that threatens it’s overall functioning capacity.

Project objective.

The reason I want to do this project is that I feel there is a better way of utilizing Mathematica to create real-world solutions for security cyber threats on a network. In doing so I can demonstrate the flexibility in creating our own tools to analyze almost anything that can be tagged as a threat on the network through the use of Mathematica. Also, many tools on the market today are not flexible enough to allow users or network administrators the ability to tweak the tools’ functionality to catch the most dangerous threats to their network. At the end of the day you really need a tool that will be able to adapt as the network changes and types of threats change as well.

Any threat analysis system requires you to conduct extensive of research to acquire knowledge on programming. Both time and resources can be minimized by using Mathematica, for pre designed functions have been already defined for the user. Other attractive features of Mathematica consist of the visual representation of data in the form of bar graphs and sorting capabilities, which in turn make it easier to analyze and store the outputs of findings for later use. In addition, users also have the option to print out the findings for future reference in a binder or audit log book.

The reason for choosing Mathematica is that it allows for the efficient analyzation of larger data sets with convenient predefined functions. These functions are very powerful and usually require no more than one liner code snippets. Also, at the University at Buffalo they currently use Mathematica for data mining. I feel that this tool can be leveraged much easier than other tools that are open source. Also, mathematical models and other types of data structures can be implemented through this program for analysis of the network traffic.

The reason for choosing GitHub is that it allows for versioning and data backup in the event of accidental file change. Any data imported and stored on GitHub can then be accessed from any operating system type. I plan on using GitHub with GitHub Desktop because I will need a way to transfer the CSV file and the pcap file to another PC that contains Mathematica. The other reason for using GitHub is that it will allow me to incorporate the data for other projects and share it with others. GitHub also has a nifty feature which allows users to view the data without actually manipulating or altering it.

**LIT REVIEW**

Similar projects that relate to the current one.

There are currently no RIT projects related to this project type. On the stackexchange forum there has been only the occasional mention of using Mathematica to view pcap files, but nothing else. This project is unique and original in nature.

**METHODOLOGY**

Steps that will be taken to show completion of the project.

What I plan on doing for this project is to monitor RIT's network with Wireshark. Since there is a policy in place related to privacy I will be monitoring a Pod in the network lab after successfully securing consent to monitor the network’s activity. The activity will be monitored for about an hour or two depending on the volunteers’ availability. After monitoring the network with Wireshark I will then attempt to save the contents to a CSV file. This data will be pushed to GitHub along with the pcap file for storage. GitHub will be utilized in order to version and keep track of any changes made to the files. Next, I will use Microsoft Access to import the data from the CSV Wireshark capture file. After completing these steps, I will use Mathematica to analyze the files for security threats (Netflix, Bit Torrent, Bitcoin, and others). In doing so I will have to program in a notebook file within Mathematica. In theory, I will be using Mathematica as a tool to make it easier for security specialists to analyze the network for security threats, as shown in **Figure 1**.

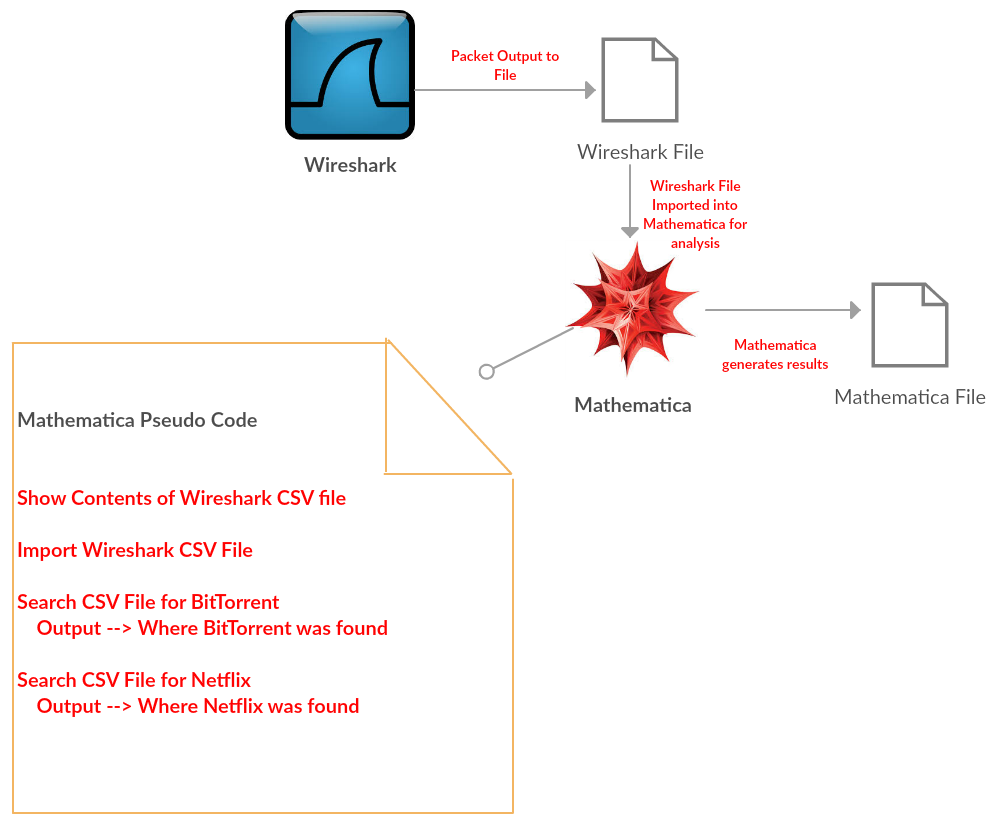


Figure 1: Mathematica Wireshark

One of the most difficult tasks involved in this multistep project is to organize the packet types into a bar graph. This will likely require the writing of a map reduce program or some other function to reduce the output of the imported files. Then after reducing the contents I will need to organize and sort the various packet types and perform a count on each packet type. As for the Bit Torrent detection a search procedure will need to be performed to sift through the data and output the lines of where the word “BitTorrent” is displayed. The same concept will be applied in searching for the words “Netflix” and “Bitcoin.”

**Important Functions for Mathematica:**

* Being able to detect Bit Torrent packets
* Being able to detect Bitcoin packets
* Being able to detect Netflix packets
* Organizing packet types into a Bar graph (UDP, TCP, etc.)
* Able to import all CSV files in the notebook folder to Mathematica
* Able to import Microsoft Access Database into Mathematica and perform previously mentioned bullet points

**DELIVERABLES**

Besides a project report, what else will be used to show the results.

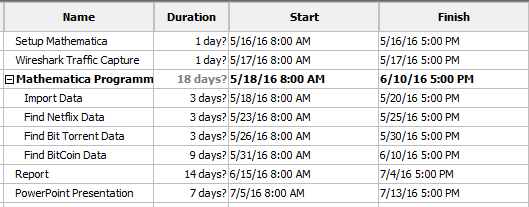
After completing this project, I plan on producing a project report in conjunction with a PowerPoint presentation. The PowerPoint presentation will provide an in-depth and visual analysis of the project, including a demo. The demo will demonstrate how Mathematica can be used to detect possible security threats. If time allows, I will try to demonstrate the ability to detect all three security threats that were explored in this project. In addition, I will also show how I was able to save the Wireshark capture to a CSV file and import it into Microsoft Access, as well as how GitHub was utilized to transfer the files between various operating systems.

**PROJECT PLAN**

Meta-analysis of the project, including the timeline and perceived level of difficulty for each step.

|  |  |  |
| --- | --- | --- |
| Wireshark Network Capture | Hard | * Need to determine where to capture to get the most accurate results. * Need to ensure that data is captured without any anomalies. |
| Wolfram Mathematica | Hard | * Need to determine how to import all the CSV files that reside in the notebook folder. * Determine how to map reduce the data to insert the various protocol types into a bar graph. |

Schedule



**Sources**

CUFP 2013. (n.d.). *CUFP 2013: Paul-Jean Letourneau: Programming Map/Reduce in Mathematica*. Retrieved from https://www.youtube.com/watch?v=pSoe-cBb\_Zg

How to Identify Network Abuse with Wireshark. (n.d.). Retrieved April 3, 2016, from http://www.howtogeek.com/107945/how-to-identify-network-abuse-with-wireshark/

osx - Access wi-fi raw data - Mathematica Stack Exchange. (n.d.). Retrieved April 3, 2016, from http://mathematica.stackexchange.com/questions/49325/access-wi-fi-raw-data

SamKearDotCom. (2013, May 27). Analyzing Bitcoin Network Traffic Using Wireshark. Retrieved from https://www.samkear.com/networking/analyzing-bitcoin-network-traffic-wireshark

Wireshark · Display Filter Reference: Bitcoin protocol. (n.d.). Retrieved April 3, 2016, from https://www.wireshark.org/docs/dfref/b/bitcoin.html