10-18-16 CSE 4243/6243 Homework 3

Instructions:

For this homework assignment, you will implement a simple network intrusion detection system. Your system will **detect**, **log**, and **alert** on:

- Any traffic to or from blacklisted IP addresses
- DNS requests for blacklisted domain names
- Unencrypted web traffic to URLs containing a set of blacklisted strings
    - Only in request URLs. These strings are fine anywhere else in a payload.
- TCP or UDP payloads containing any of a set of simple signatures
    - Signatures can be any arbitrary byte sequence (not limited to just text)
    - May include null bytes within the sequence
- Network port scanning activity

Your system should run from the command line, allowing a user to provide (as command line arguments) a **PCAP** format packet dump or a network interface that will be monitored in promiscuous mode. Other command line arguments should be used to specify files that contain the detection signatures and blacklists described above (they may be in separate files, or combined). Configuration files should be human-readable and editable, which may require some care when considering how the TCP/UDP payload signatures are specified (as they are not limited to plaintext). Network port scan detection should not require a configuration file (unless you wish to have tunable parameters for that detection).

Your system will **detect** traffic that matches the circumstances described above. Every match should be **logged** to a text format log file (the name should be provided as a command line argument). The log file should contain timestamps for each entry and a description of what signature was detected and any other relevant information. **Alerts** should be provided in the visual output of your program for matches as well, but to prevent the alerts from being overwhelming, alerts should not repeat the same output any more frequently than every five minutes (while the logs contain a more complete set of output).

Your system may be developed and run on any operating system you choose, using any programming language you are comfortable with. You may choose any library you wish for packet capture and manipulation with the exception of any libraries designed specifically for embedding or implementing intrusion detection functionality. You may also need to develop small programs for generating traffic used in testing.

Your deliverables should be in a **zip** format file containing:

- Source code for your system, and any programs developed to test it.
- Compiled binaries for your system (if you are using a compiled language)
- Installation and complete usage documentation
- A report describing
    - Your decision process on language/library use
    - The design of your system
    - Any difficulties encountered in developing your system
    - A description of any limitations of your system, with regards to its ability to detect or perform
    - Testing – **Describe** and **Document** the process and results from testing all of the functionality of your system. Also, test and describe your system's ability to handle a high speed/volume of traffic and document any limitations it may have in these situations.

You may find it convenient to use virtualization to develop and test your system. This would allow you to send traffic between your host operating system and a virtual machine, or between two or more virtual machines. 30-day trial versions of VMware are available on their site. Virtual Box is a free alternative, though the networking functionality is not as straightforward.

As stated in the syllabus for all assignments in this course, this assignment is strictly for **individual work only**. The purpose of this assignment is to demonstrate **your** capability of developing a system that meets the requirements. You are not to discuss or share any aspect of this assignment with other students, including verbal discussion, code or resources.

You are expected to cite all sources used in the creation of your deliverables. This includes library documentation and any other online sources uses. **All code is expected to be of your own creation**. For full credit on the assignment, you will limit yourself to primary sources of documentation (official language and library documentation, as well as the official documentation of libraries you utilize). Submissions that use other sources (for example: blogs, tutorials, "question" sites like StackExchange) will not receive full credit. If you utilize a source without citing it, that constitutes plagiarism and will be reported through the Honor Code office.

Grading

80 – System

- 10 - Detection of IP addresses
- 10 - Detection of DNS requests
- 15 - Detection of URL matches
- 15 - Detection of TCP/UDP payload signatures
- 15 - Detection of port scans
- 15 – Remainder of described functionality (interface, alerting, logging, etc.)

20 – Report

- 10 – Report requirements
- 10 – Testing

Note that good reporting can assist me in determining correct functionality and in assigning partial credit. If I can't get something to work, and there's no documentation of it ever having worked, you will lose points.

Your code is expected to be well-formatted, commented, and easy for a peer to follow. Your report must exhibit good technical writing style, grammar, spelling, and formatting. Points will be deducted from either section independently from the above grading scheme for issues with code or reporting style.