Assignment-1
DFSC 2316
DF and IA Fundamentals – II
100 points

1. In this question, we tackle the issue of malware outbreak on ***QuickDryingCement*** Corporation's server in the context of *Network/Digital Forensics*. Your forensics firm ***OnlyTheBest*** has been contracted to collect and analyze the logs from the server's iptables/firewalls. ***Joe Somebody***, a respected colleague in your firm, has collected the logs and given it to your team for further research and analysis. His preliminary findings have led to the conclusion that the corporation's server may have been infected with one or more malware, namely ***WinCrash***, ***Scarab*** and/or the ***Doly Trojan***. Your task is to examine the given log in the archive, parse the log structure to its constituent parts to completely comprehend its meaning and assess whether or not there are indications of these malicious viruses/worms on the machine. Please prepare some simple visuals (Excel plots) and make appropriate recommendations to ***QuickDryingCement's*** board. You will need the ***NetworkTraffic.log*** file in the given archive. You system administrator has strongly recommended that (i.e., it is imperative) that you conduct your analysis on a Linux system and use appropriate tools on your Kali Linux virtual machine under Virtual Box. You are welcome to write your own scripts if you prefer not to use existing tools. The log file contains network traffic for the month of May 2015. Conduct your analysis for every day of the month of May and enter your counts into a spreadsheet (Date, WinCrash count, Scarab count, Doly Trojan count) and graphically display each count by date. Include your graph, interpretation of its meaning and recommendations for further action to ***QuickDryingCement's*** system administrator. In particular, annotate any anomalies that you may find (based on your investigation) in your report.

   Learning Objectives, hints and motivation for the problem:

   a. This exercise should pique your interest towards network forensics tools in Kali Linux.
   b. First, you should use appropriate archiving tools to extract the contents of the given archive. (Hint: ***`man tar`***)
   c. Next, you should try to analyze the contents of the extracted ***NetworkTraffic.log*** file and "understand" its structure. Effectively, I am asking you to parse the structure of the file before proceeding to the next step.
   d. Having gathered a reasonable understanding of the ***NetworkTraffic.log*** file, proceed to "pattern match" appropriately chosen patterns to extract portions of the ASCII text log file, using any tool that you want. (Hint: ***`man grep`***)
   e. Now, you should go ahead and figure out which network ports are used by the malware ***WinCrash***, ***Scarab*** and the ***Doly Trojan.*** (Hint: ***Google***!)
   f. Now that you know which ports are used by the above mentioned malware, go ahead and "count" the number of occurrences using a reliable tool in Linux (Hint: ***`man wc`***)
   g. Now, you are ready to plot the results on Excel or similar plotting tool. Summary: This problem hopes to assist you in learning some elementary "log" analysis using popular Unix tools ***grep*** and ***wc***.

As a complete example, here is how I would do these tasks on Windows. I recommend performing these tasks in Windows and then proceeding to solve the problem under Linux.

The Windows ***find*** command searches for strings in a file. Its general syntax is:

find "string" file

It's full syntax can be examined by typing ***help find*** in a command/cmd/terminal/console window. Its default behavior is to display lines from the file that contain the string but it also has a ***/c*** switch to just display the number of lines containing the string.

I would think the most prudent, simple first step to accomplish is to separate the log into "daily" log files and this will be done based on the date. For example, the command:

find "May 1 " NetworkTraffic.log > May01.log

will copy the May 1st records to the file ***May01.log***. Note that there is a space after the 1 in the string (this avoids also matching 11, 12, 13, 14, 15, 15, 17, 18 and 19). Perform similar steps to produce daily logs for the remainder of ***May***.

The next step is to determine which port is most commonly associated with a particular malware. For instance, the port commonly associated with the MyDoom virus is 3127. Try to decipher how many log records contain the ports of interest. For example:

find /c "DPT=3127" May01.log

will display the count of the number of log records with a destination port of 3127 on May 1st. Repeat the above steps now on Linux for all pieces of malware that you are interested in identifying and complete the problem.

Instructions: Please answer the question in as much detail as possible. In all cases, demonstration of work using clear and readable screenshots on your machines is expected. Answer all the questions very carefully and in a comprehensive manner. Incomplete/vague/partial answers that demonstrate no work will fetch little or no points.