Nolan Aubuchon

Dr. Wahjudi

CS 430

April 21, 2012

Project 5 – Napalm

For this project, we had to take a log of network traffic and parse it using Python. There is a total of four parts to the project. For part one we had to find a list of all the protocols being used and find the percentage of the total for each of them. For part two we had to find a list of IP addresses that tried to connect to an entered IP address. For part three we had to find a list of FTP servers, their valid usernames, and their valid passwords. For part four we had to find all the MAC addresses and find their corresponding IP addresses. The method I used is using **regular expressions** to parse the text.

|  |  |
| --- | --- |
|  | Regular Expression |
| Part 1 | “->\s+\S+\s+(\w+)” |
| Part 2 | “\s(\d+\.\d+\.\d+\.\d+)\s->\s(\d+\.\d+\.\d+\.\d+)\s” |
| Part 3a | “FTP\s+\d+\s+Request:\s+USER\s+(\S+)” |
| Part 3b | “FTP\s+\d+\s+Request:\s+PASS\s+(\S+)” |
| Part 3c | “FTP\s+\d+\s+Response:\s+(\d+)” |
| Part 3d | “->\s+(\S+)\sFTP” |
| Part 4 | “(\w\w:\w\w:\w\w:\w\w:\w\w:\w\w)\s\*->\s\*(\w\w:\w\w:\w\w:\w\w:\w\w:\w\w).\*Who has (\S\*)\?.\*Tell (\S\*)” |

For example, the first line of the web traffic dump *“ 1 0.000000 31.205.0.166 -> 10.200.59.77 TCP 54 51244 > ssh [ACK] Seq=1 Ack=1 Win=253 Len=0*” Matches the regular expression *“->\s+\S+\s+(\w+)”* . The parentheses match the “TCP” part of the line, as the goal of part 1 is to find the different protocols being used in the web traffic. For the second part, I use regular expressions to search for groups of digits separated by 3 periods; these groups of digits make up the IP addresses to parse, e.g. “128.128.128.128 -> 255.255.255.255”. For part 3, we’re looking for someone trying to enter their username, followed by their password, followed by the HTTP reply code ’230’, which means the username and password were valid. For part 4, were looking for 6 pairs of word characters (zero through nine, A through Z, and underscore) separated by colons, e.g. “00:00:00:00:00:00 -> ff:ff:ff:ff:ff:ff” which are followed by their IP addresses in the same line of web log, e.g. “Who has 128.128.128.128? Tell 255.255.255.255” In this case the IP address 255.255.255.255 has the MAC address 00:00:00:00:00:00, whereas the IP address 128.128.128.128 has the MAC address ff:ff:ff:ff:ff:ff.