Incident Response Technologies
Homework 4: Splunk

Cliff C. Zou

My research group has a web server set up on a Linux machine, and I have downloaded a fraction of the Apache webserver's log files as 'apache2.zip', which contains two log files: access.log and error.log, and you can download it from the assignment webpage. You need to import this apache2.zip data into your Splunk software and analyze it in order to answer the following questions. **Note: You must provide answers based on Splunk usage, not based on any other analysis tools**.

Question 1 (20 points): How many events are in 'access.log' and in 'error.log', respectively? Your answer should include one or several screenshot images of these numbers.

Question 2 (20 points): How many events happened between Oct. 6th 4:20:00pm, 2017 to Oct. 6th 4:21:00pm, 2017? Please also show the **pie chart** graph, and the statistics of the '**status**' field for this time range events. You need to use screenshot image to show the pie chart graph and the statistics of the status field.

Question 3 (30 points): When checking the 'access.log', the majority of client requests showed that they were sent from 'Mozilla' browsers. There are multiple versions of Mozilla browsers used by web clients. We are interested in the browser versions distribution, but the default fields in Splunk do not have such a field extracted for analysis.

In this question, you are required to extract the browser version field in Splunk, then we can rely on Splunk to give us the browser version statistics. For example, for the following event:

10/6/17 4:27:51.000 PM	10.171.14.38 - - [06/Oct/2017:16:27:51 -0400] "POST /projects/icaptcha/main.php/GET/webstyle.css/webstyle.css/query.php HTTP/1.1" 200 4189 "-" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0)"

Its browser version is '4.0'. Please extract this value with the field name of 'Browser_version'.

For events happened during the entire day of Oct. 6th, 2017, how many different browser versions exist? What is the number of events for each value of the browser version? You need to use screenshot image to show the results.

Question 4 (30 points): In class I have demonstrated how to show geolocation graph based on client IP addresses, and how to add generated graph to dashboard. In this question, you are required to:

(1) Show the search term you used to obtain the geolocation graph based on clientip field extracted by Splunk on the entire data log zip file (on 'all time' time range). The geostats command should be used with 'count by status' setting.
(2) Show the screenshot image of the generate 'Cluster map' of the geolocation graph.
(3) Add this geolocation graph into a newly created dashboard titled 'CIS6395-hw4', with the panel title 'client IP geolocation'. Show the screenshot image of the created dashboard, including the dashboard title and the panel title.