Assignment 2

Instructions (READ CAREFULLY!!!)

1. Follow the directory format provided. Assignments that violate the directory format will be considered as not having been submitted. The directory can be downloaded here. **Remember to change the name of root folder to your student ID and to re-zip into a zip folder named your student ID!**

2. Complete this assignment in one of either Python 3.7, C, C++, or Java. If using Java, do not use packages as they would violate the directory format and complicate command-line compilation. Also, if using Java, ensure that your file paths for input and output files are not specific for your machine; submissions with this problem will be awarded 0. You should name your entry file main.¡ext¿ where ¡ext¿ is the appropriate file extension for your choosen language; for example, if you used Python, your file should be named main.py

3. **REMEMBER THAT YOU ARE BEING MARKED ON TEST CASES (PER SLIDE# 22 OF LECTURE 1). HENCE CODE THAT DOES NOT COMPILE, PRODUCE THE REQUIRED OUTPUT FILES, CRASHES DURING EXECUTION, ETC... WILL BE AWARDED A MARK OF 0**

4. Ensure that you edit the JSON file in the root of the directory with your name, Student ID, programming language used, and Email address.

5. Sign and attach either a soft copy or clear image of the plagiarism declaration with your submissions. There is a folder in the directory for you to store this declaration. You are allowed to discuss the assignment with your classmates; however, your code and write-up MUST be your own. You are free to use the sample code provided on the course Github repo.

6. After unzipping the directory structure, edit with the confines of that directory structure and then places it back into a .zip file with your

Student ID. Other compression formats such as .rar or .7zip, or archives not named your Student ID will be rejected.

7. After completion of your assignment, fill out **this form** and email a copy of the assignemnt to irahamancourses@gmail.com. Your email should have the subject COMP2611 A2 1234 where 1234 is your student ID number. YOU MUST DOWNLOAD THIS PDF TO CLICK THE LINK TO THE FORM!

8. You are free to use any in-built data type in Python, C++, or Java (sets, dictionaroes, vectors, etc....). However, you must store your data in Binary Search Tree (BST) - none of these provide a BST type.

9. - Draft Date: 31st October 2019
   - Final Hard deadline: 7th November 2019

Part A - Determining Access

Every machine on the Internet has an IP address. When a computer, called the client, requests data from another computer, called the server, the server logs the date and time of the request (typically as a UNIX timestamp) and the client's IP address.

A server facilitating identity theft has been confiscated by the cyber-crime division. The division is aware of the approximate time and approximate ranges of time of certain important transactions. Using the data in the server logs, you are tasked with writing code that can efficiently retrieve IP addresses that accessed the server during or at those time periods.

Assuming that each UNIX timestamp in the log is unique (i.e. no two IP addresses access the server at the exact same time), you can store each access record in a Binary Search Tree using the timestamp as the key.

There are three forms of queries that your programme needs to allow:

- LTE x - retrieves the IP address that accessed the server at time x or last access before x
- GTE x - retrieves the IP address that accessed the server at time x or the first access after x
- BTN x y - retrieves all IP addresses that accessed the server between time x (inclusive) and time y (inclusive)

You will provided with the following input files:

- log.txt - each line contains a UNIX timestamp, followed by a space, and then an IPv4 address
- queries.txt - each line contains a query of the forms specified above

Each line in your output file - solution.txt - should contain the answer to the corresponding query in queries.txt. Each IPv4 address in an answer should be separated by a single whitespace. If no IP address is found for a query, the corresponding line should contain NONE

NB:

A UNIX timestamp is an integer with less than 11 digits. Hence, if using C or C++, unsigned long long int should be sufficient to store them. Each IPv4 address has a maximum of 15 characters.