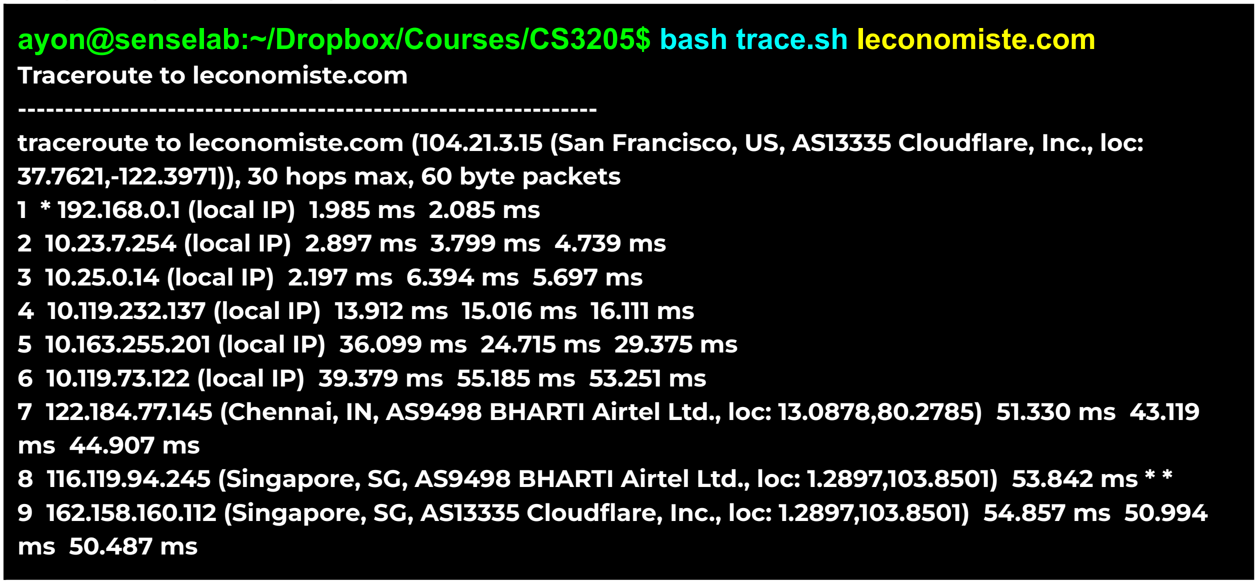
CS3205 (Semester: Holi 2025) Programming Assignment 3

Submission Deadline: *Sunday, 4th May, 2025* Total: *35 Marks*

Exercise 1. [20 marks, Internet Measurement Analytics]. Explore traceroute tool to reach the following servers located in the following African countries. Use this script to generate the routing trace along with RTTs and geolocations of the routers in the path. We use the geolocation service provided by https://ipinfo.io, we also output the approximate location of the router, as reported by the API, in terms of latitude and longitude.

1.	Angola	paratus.ao
2.	Botswana	weekendpost.co.bw
3.	Burkina Faso	onatel.bf
4.	Cameroon	camtel.cm
5.	DR Congo	vodacom.cd
6.	Libya	ltt.ly
7.	Madagascar	telma.mg
8.	Mali	afribone.com
9.	Morocco	orange.ma
10.	Namibia	namhost.com
11.	Rwanda	ricta.org.rw
12.	Senegal	sonatel.sn
13.	Uganda	ucc.co.ug
14.	Zambia	zamnet.zm
15.	Zimbabwe	zol.co.zw

Sample output of the script:



```
ayon@senselab:~/Dropbox/Courses/CS3205$ bash trace.sh leconomiste.com
Traceroute to leconomiste.com
-----------------------------------------------------------
traceroute to leconomiste.com (104.21.3.15 (San Francisco, US, AS13335 Cloudflare, Inc., loc:
37.7621,-122.3971)), 30 hops max, 60 byte packets
1  * 192.168.0.1 (local IP)  1.985 ms  2.085 ms
2  10.23.7.254 (local IP)  2.897 ms  3.799 ms  4.739 ms
3  10.25.0.14 (local IP)  2.197 ms  6.394 ms  5.697 ms
4  10.119.232.137 (local IP)  13.912 ms  15.016 ms  16.111 ms
5  10.163.255.201 (local IP)  36.099 ms  24.715 ms  29.375 ms
6  10.119.73.122 (local IP)  39.379 ms  55.185 ms  53.251 ms
7  122.184.77.145 (Chennai, IN, AS9498 BHARTI Airtel Ltd., loc: 13.0878,80.2785)  51.330 ms  43.119
ms  44.907 ms
8  116.119.94.245 (Singapore, SG, AS9498 BHARTI Airtel Ltd., loc: 1.2897,103.8501)  53.842 ms * *
9  162.158.160.112 (Singapore, SG, AS13335 Cloudflare, Inc., loc: 1.2897,103.8501)  54.857 ms  50.994
ms  50.487 ms
```

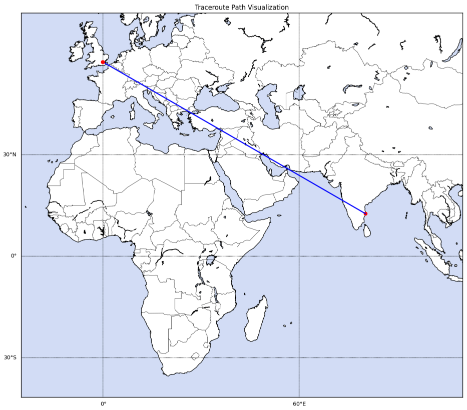
```
10  172.69.117.3 (Singapore, SG, AS13335 Cloudflare, Inc., loc: 1.2897,103.8501)  42.621 ms
162.158.160.137 (Singapore, SG, AS13335 Cloudflare, Inc., loc: 1.2897,103.8501)  63.891 ms
162.158.160.141 (Singapore, SG, AS13335 Cloudflare, Inc., loc: 1.2897,103.8501)  42.793 ms
11  104.21.3.15 (San Francisco, US, AS13335 Cloudflare, Inc., loc: 37.7621,-122.3971)  41.108 ms  52.556
ms  43.059 ms
```

Your task is to run the script for each country, **10 times** (i.e., you can have ~30 RTTs per router) and log all outputs in a single txt file and analyze it. You can automate the trace collection if needed. Use an iPython/Jupyter notebook.

a. How many entries are there totally in all of your measurement traces? Ignore all hops where the router identities are not revealed. If some RTTs are invalid (e.g., see the 8th hop, **53.842 ms * ***) only consider the valid ones.

b. How many unique router IPs can be seen? For each router report the average RTT from source (Chennai) and the standard deviation of RTT. What are the top 5 routers with the highest standard deviation?

c. On a world map show the location of all the routers (vertices) and show their connectivity with a straight line (edge). Any two routers whose reported locations are within 50 Kms of each other can be treated as a single vertex. The link or edge line should be colour coded indexed with a colorbar where the colour value indicates the average latency of the link. If there are intermediate routers that are not reported or don't respond to ICMP ignore them. For example, if you see an average RTT of 50 ms at hop 10 and an average RTT of 75 ms at hop 13, where hops 11 and 12 are unreported, you may assume that the edge weight from Hop 10 to 13 is 75-50 = 25 ms. (*Use this sample python snippet to draw the base world map. Example output of the snippet is shown on the right, two points connecting Chennai and London with a straight line).*



d. Draw a scatter plot with x-axis indicating physical distance between the routers and y-axis indicating the latency associated with that edge. Fit a trend line and compute the correlation coefficient. (*Use haversine distance for x-axis, don't take Euclidean distance between [lat,lon] pairs! FYI, lat and lon indicate angles*)

e. For some hops, you may see that the packets travel through three different routers for three queries. Example, see the 10th hop in the provided trace. This is called per packet load balancing. How many such cases are observed in total? At what locations and under which ISP?

f. Do you see any specific router(s), outside India, that are common to most of the Internet paths for the listed destinations?

g. **(optional)** Explore the tool called paris-traceroute and understand the advantages that it has over vanilla traceroute. (https://paris-traceroute.net/, The Power of Paris Traceroute for Modern Load-Balanced Networks | Kentik Blog).

Exercise 2. [15 marks, Routing Algorithms]. Consider only the routers found in the previous exercise as vertices of a graph and consider all edges (complete graph). The locations of the vertices are as recorded. Assign them IDs from 1 to N. The weight of the edge is proportional to the physical (haversine) distance between the routers. Take a variable called threshold, T, within the range 0 and 1. Randomly drop T fraction of the edges. USe reasonable values, don't have a too dense or a too sparse graph.

a. Implement either (i) or (ii):
   i. The Link State Routing protocol. Count the total number of LSA messages put in the network during the flooding stage. What is the number of entries in the Link State Database? How many (time slots) rounds were necessary to have the database populated in all routers?
   ii. The Distance Vector Routing protocol. Count the total number of distance vectors exchanged in the network. How many rounds were necessary to have the updated distance vectors in all routers?

   It is important to give a full explanation of your implementation strategy (not just code comment one liners) to obtain full credits.

b. Implement a function that takes the router ID and prints the forwarding table.
c. Implement a function that takes the router source and destination IDs and provides the routing path.

Submission Instructions: Create a folder called CS3205_Assignment3_roll1_roll2. Have two subfolders: ex1 and ex2. In each folder have the relevant notebooks and traces. Compress the root folder as a gzipped file and submit CS3205_Assignment3_roll1_roll2.gzip

Late Submission. The deadline for submitting this assignment is 11:00 pm, 4th May. After this there will be a late penalty of 25% for each day for the next four days. You are not required to submit the assignment solutions beyond 8th May.

Viva. The viva will be conducted online if felt necessary. Each group will be individually contacted in case viva is required. This viva will be conducted by myself - TAs will not be involved.