Visualizing the Internet Topology
Assignment 1: CS3530

Introduction

To visualize the Internet topology, the 'traceroute' command was used. We selected 5 sources and 10 destinations to perform the traceroutes. The traceroute command performs a series of hops between the source and any chosen destination, and the RTTs for every hop (source to intermediate router) are displayed for 3 packets. The raw data we got from the traceroute output was then parsed and processed to create a list of every IP address and corresponding AS number.

Data Acquisition:

First, the sources and destinations had to be chosen. Two sources - from IITH LAN and mobile hotspot - were done locally by implementing the traceroute command on the terminal. The 3 other sources - Paris, New York and Seoul were chosen from Looking Glass websites.

The destinations - first, some popular websites were chosen such that none of them were hosted by the same ISP. We chose 10 destinations from 3 countries - US, UK, Japan. After finalizing the sources and destinations, and making sure that the destinations are being reached by our traceroute, we directed the traceroute outputs to 10 different textfiles. Every source's traceroute was appended in the corresponding .txt file of the destination.

Details of the destinations are as follows-

Website	Location of server
www.columbia.edu	New York, USA
www.cmu.edu	Pittsburgh, USA
www.discord.com	San Francisco, USA
www.craigslist.org	Virginia, USA
robert-schumann.com	Osaka, JP
pc.watch.impress.co.jp	Tokyo, JP

tripmall.online	Osaka, JP
cam.ac.uk	Cambridge, UK
www.virginmedia.com	London, UK
www.talktalkgroup.com	York, UK

Data Processing:

Now we have the traceroute results for 10 destinations in separate files with outputs from each of the 5 sources appended to the same destination file. The traceroute results contain the IP addresses of intermediate routers and round trip time delay for each router in the path, and so a python script was written to extract this information and store it in an edge list format.

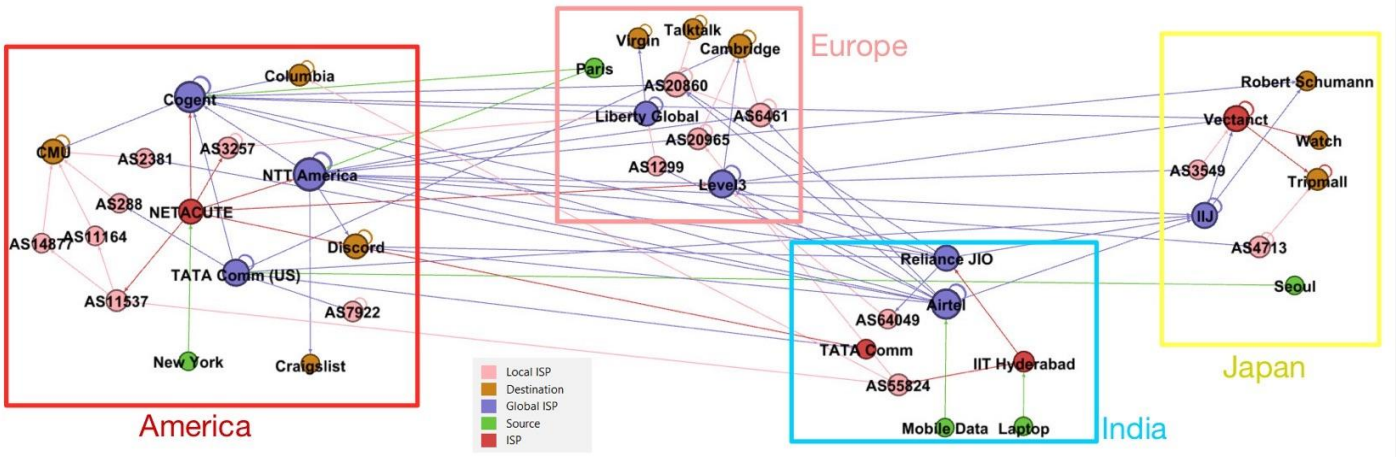
Edge list essentially stores information of the source and destination router, in our case we store the previous and the current IP address from the traceroute output (since they form an edge in the network) and the difference in time delay from the current traceroute and previous traceroute result (which signifies the time delay in transmission from one node to the other).

In the same script using the IPWhois package in python, we extracted the AS numbers of each IP address from the edge list. It was observed that multiple IP addresses were using the same AS numbers and some IP addresses in the path of the network were private and hence did not have an AS number. So these private addresses were removed from the path and combined the IP addresses according to the range provided by that AS number.

To extract the geolocations of each router in the path, another script was written that sends a request to this website with an IP address and it returns details about that router like City, Country, Latitude and Longitude.

Visualization:

After extracting the edge list for each source, the network was plotted in Gephi, a graph plotting software. Every node is an AS number, and every directed edge is a hop from one router to the next.



First two sources were the laptop over IITH lan and with Airtel Mobile data. For the third source we used New York from the website Perfops, Seoul from TATA looking glass and Paris from NTT looking glass. The choice of looking glass/website/network provider explains the first hop from the source.

AS Numbers:

Each AS number have been assigned their own IP address ranges (multiple are possible). In the following table we list every AS number we encountered and the IP address range we came under for the respective AS number.

The AS number lookup table is appended to the end of this document.

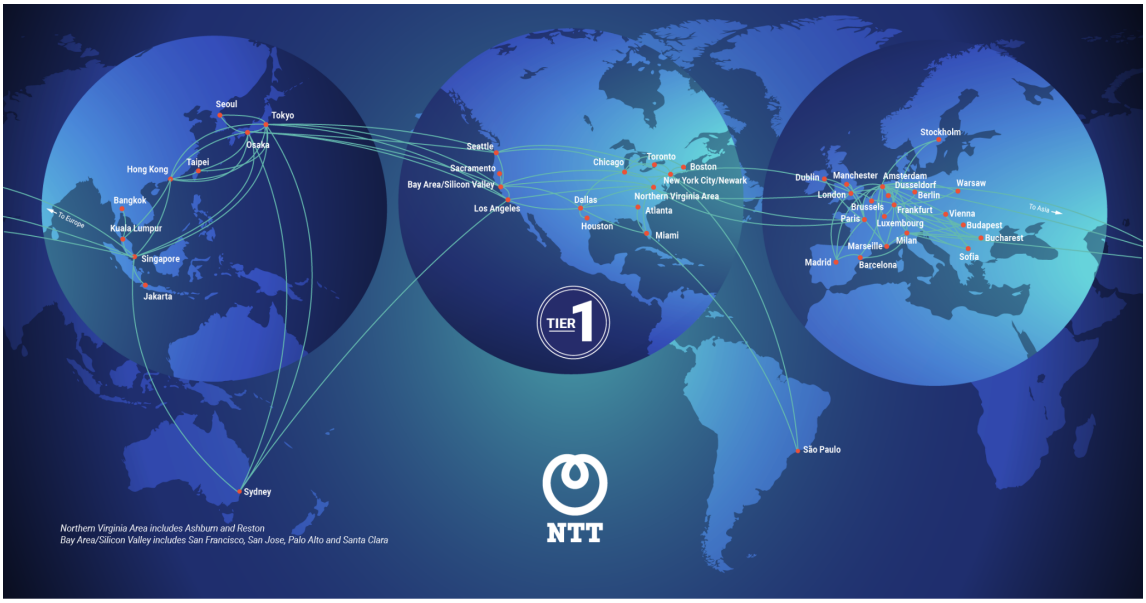
Findings about Network Topology:

We learned many interesting facts about the Internet while doing this assignment.

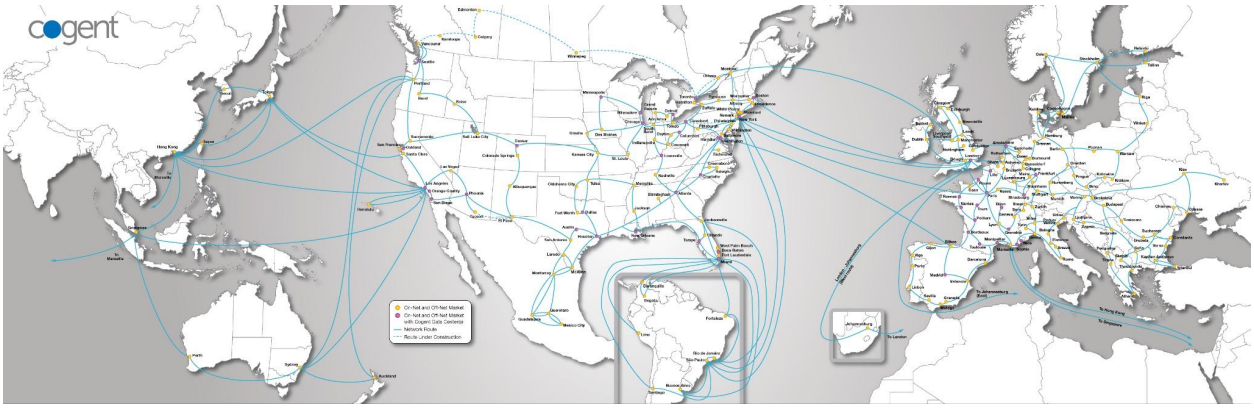
1. When using Looking Glass, say NTT's looking glass, traceroutes from different source routers to the same destination did not have too much variation among them. After the first 1-2 hops, the route being followed was more or less the same. And, all of these routers come under the same AS number. After looking at the complete network of NTT we realized that the Looking Glass will route internally first, via the NTT network, and then route to the actual destination from the

closest router in NTT. So, the source AS number would be the same for different routers - will virtually look the same. For this reason, we tried to select other networks' Looking Glass for each source.

The two major players in ISP providers are NTT and Cogent, here are their networks:



NTTs Network



Cogent's Network

2. Some big providers like Tata Communications India look small in our representation even though they are huge players in the market which can be attributed to the selection of sources and destinations. Vice-versa is also true that smaller ISPs are depicted as having more interconnections.

3. While searching for good destination IP addresses to route to, we realized that some websites did not have a constant IP address. So, a new IP address under the same AS number was allocated every couple of minutes/seconds. We avoided those websites for the purpose of this assignment.

4. When looking up the AS numbers from the IP addresses, we found many private networks for which the AS number was not provided. There were many private IP addresses in the vicinity of IITH, but sometimes a few private routers in the middle as well. We skipped over these in the visualization for neatness and simplicity.

5. We used the '-I' option for executing the traceroute command. This option was to ensure ICMP packets were being sent, instead of UDP. Many a time, traceroute with UDP packets was not reaching the destination appropriately. Tracert in Powershell, however, sends ICMP packets by default.

Group Members:

Anshul Gupta - EE20BTECH11004
Padmini Palivela - EE20BTECH11038
Pushkal Mishra - EE20BTECH11042

AS Number	IP Range	City	Country Code	Description
9	128.2.0.0 - 128.2.255.255	Pittsburgh	US	Carnegie Mellon University
14	128.59.0.0 - 128.59.255.255	New York	US	Columbia University
174	38.0.0.0 - 38.255.255.255	Pittsburgh	US	PSINet, Inc.
	66.28.0.0 - 66.28.255.255	Washington	US	Cogent Communications
	130.117.0.0 - 130.117.255.255			PSINet, Inc.
	149.11.0.0 - 149.11.255.255	Camden	GB	PSINet, Inc.
	154.54.0.0 - 154.54.255.255		US	PSINet, Inc.
786	128.232.0.0 - 128.232.255.255	Cambridge	GB	University of Cambridge Computer Laboratory
	131.111.0.0 - 131.111.255.255	Cambridge	GB	University of Cambridge
	146.97.0.0 - 146.97.255.255		GB	Jisc Services Limited
	193.60.80.0 - 193.60.95.255	Cambridge	GB	University of Cambridge
1299	62.115.128.0 - 62.115.143.255			Arelion Sweden AB
2381	216.56.0.0 - 216.56.255.255	Waukesha	US	WiscNet
2497	58.138.0.0 - 58.138.127.255		JP	Internet Initiative Japan Inc. Iidabashi Grand Bloom, 2-10-2 Fujimi, Chiyoda-ku, Tokyo, 102-0071 Japan
	202.232.0.0 - 202.235.255.255		JP	Japan Network Information Center
	210.128.0.0 - 210.135.255.255		JP	Japan Network Information Center
	210.136.0.0 - 210.143.255.255		JP	Japan Network Information Center
2519	163.139.0.0 - 163.141.255.255		JP	Japan Network Information Center
	222.230.0.0 - 222.230.255.255		JP	ARTERIA Networks Corporation
2828	66.2.0.0 - 66.3.255.255	Cleveland	US	Verizon Business
2914	61.112.0.0 - 61.127.255.255		JP	Japan Network Information Center
	120.88.48.0 - 120.88.63.255	Tokyo	JP	NTT Ltd Japan Corporation
	128.241.0.0 - 128.241.255.255	Sherman Oaks	US	NTT America, Inc.
	129.250.0.0 - 129.250.255.255		US	NTT America, Inc.
	165.254.0.0 - 165.254.255.255	Santa Monica	US	NTT America, Inc.
	168.143.0.0 - 168.143.255.255		US	NTT America, Inc.
	185.84.18.0 - 185.84.18.255	Paris	FR	VERIO FR paris facility
3257	89.149.181.0 - 89.149.181.255		IE	Tinet International Network
	173.241.128.0 - 173.241.143.255	Somerville	US	Tinet
	4.0.0.0 - 4.127.255.255		US	Level 3 Parent, LLC

3356	8.244.0.0 - 8.255.255.255		US	Level 3 Parent, LLC
	212.187.216.0 - 212.187.216.255		GB	
3549	113.29.0.0 - 113.29.127.255	Chiyoda-ku	JP	Level 3 Communications, Inc. 1025 Eldorado Blvd. Broomfield, CO 80021
4694	202.216.0.0 - 202.219.255.255	Osaka	JP	Japan Network Information Center
4713	122.1.0.0 - 122.1.255.255		JP	NTT Communications Corporation
	210.232.0.0 - 210.235.255.255		JP	Japan Network Information Center
4755	115.112.0.0 - 115.119.255.255		IN	Internet Service Provider TATA Communications formerly VSNL is Leading ISP, Data and Voice Carrier in India
5089	62.252.0.0 - 62.255.255.255		GB	Virgin Media
	213.104.0.0 - 213.107.255.255	Stockbridge	GB	Virgin Media
6453	63.243.128.0 - 63.243.255.255		US	TATA COMMUNICATIONS (AMERICA) INC
	64.86.0.0 - 64.86.255.255		US	TATA COMMUNICATIONS (AMERICA) INC
	66.110.0.0 - 66.110.127.255		US	TATA COMMUNICATIONS (AMERICA) INC
	116.0.64.0 - 116.0.95.255		SG	Tata Communications (CANADA) Ltd. Global Customer Service Centre Alandi Road, Dighi, Pune
	116.0.67.0 - 116.0.67.255	Central	HK	customers access -30 and BB use
	116.0.93.0 - 116.0.93.255		SG	Customers access -30 and BB internal use
	120.29.211.0 - 120.29.211.255		JP	Customers access -30 and BB internal use
	120.29.217.0 - 120.29.217.255		IN	Customer access-30 and BB internal use
	180.87.3.128 - 180.87.3.255		IN	Customers access /31 and BB internal use
	180.87.37.0 - 180.87.37.255		IN	Customers access -30 and BB internal use
	180.87.112.0 - 180.87.112.255	Central	HK	Customer access /30 & BB usage
	180.87.181.0 - 180.87.181.255		IN	Customers access -30 and BB internal use
	209.58.0.0 - 209.58.127.255		US	TATA COMMUNICATIONS (AMERICA) INC
6461	64.124.190.0 - 64.125.255.255		US	Zayo Bandwidth
	94.31.0.0 - 94.31.63.255	Hammersmith	GB	
6830	84.116.64.0 - 84.116.207.255			Liberty Global Infrastructure
	213.46.174.0 - 213.46.174.255		NL	Chello Broadband Links and Loopbacks in the UK
7922	50.128.0.0 - 50.255.255.255	Hayward	US	Comcast Cable Communications, LLC
	68.80.0.0 - 68.87.255.255		US	Comcast Cable Communications, LLC
	96.64.0.0 - 96.124.255.255		US	Comcast Cable Communications, LLC
9371	219.94.128.0 - 219.94.255.255	Osaka	JP	SAKURA Internet Inc.

9498	116.119.0.0 - 116.119.127.0		IN	BHARTI TELESONIC INFRASTRUCTURE
	125.18.92.0 - 125.18.92.255		IN	Bharti Infotel Ltd. (BB&TS) Mumba
	182.64.0.0 - 182.79.255.255		IN	Bharti Airtel Limited
11164	198.71.44.0 - 198.71.47.255		US	Internet2
11537	163.253.0.0 - 163.253.255.255		US	Internet2
13335	141.101.67.0 - 141.101.67.255	Paris	FR	CloudFlare, Inc.
	162.158.0.0 - 162.159.255.255	Not found	Not found	Cloudflare, Inc.
	172.64.0.0 - 172.71.255.255		US	Cloudflare, Inc.
14877	162.223.16.0 - 162.223.19.255		US	First Light Fiber
20860	62.128.192.0 - 62.128.223.255		GB	
	130.180.202.0 - 130.180.202.255		GB	SERVERSPACE-NET:::Infra
20965	62.40.96.0 - 62.40.111.255		GB	IP allocation for GEANT network
	62.40.112.0 - 62.40.127.255		GB	IP allocation for GEANT network infrastructure
22414	208.82.236.0 - 208.82.239.255		US	Craigslist, Inc.
35425	5.28.62.0 - 5.28.62.255		GB	BIGV1-YRK-VMACC1
	91.223.58.0 - 91.223.58.255		GB	
36236	199.38.180.0 - 199.38.183.255	New York	US	NetActuate, Inc
55824	180.149.48.0 - 180.149.48.255		IN	NKN South Universities
	115.240.0.0 - 115.247.255.255		IN	Reliance Jio Infocomm Limited
59193	103.232.241.0 - 103.232.241.255	Champapet	IN	IIT Hyderabad
64049	49.32.0.0 - 49.47.255.255		IN	Reliance Jio Infocomm Limited
131965	103.3.0.0 - 103.3.3.255		JP	XSERVER Inc.
	120.136.8.0 - 120.136.15.255	Osaka	JP	XSERVER Inc.