CSCI4171 Assignment1

1.(a)

- The traceroute is displaying the route and measuring transit delays of packets across an internet protocol network. When you enter the traceroute command, the utility initiates the sending of a packet, including in the packet a time limit value that is designed to be exceeded by the first router that receive it, which will return a Time Exceeded message. This enables traceroute to determine the time required for the hop to the first router. Increasing the time limit value, it resends the packet so that it will reach the second router in the path to the destination, which returns another Time Exceeded message and so forth. Traceroute determines when the packet has reached the destination by including a port number that is outside the normal range. When it is received, a Port Unreachable message is returned, enabling traceroute to measure for you hop by hop.
- Traceroute syntax:
  [**-m** *Max_ttl*] [**-n**] [**-p** *Port*] [**-q** *Nqueries*] [**-r**] [**-s** *SRC_Addr*] [**-t** *TypeOfService*] [**-f** *flow*] [**-v**] [**-w** *WaitTime*] *Host* [*PacketSize*]

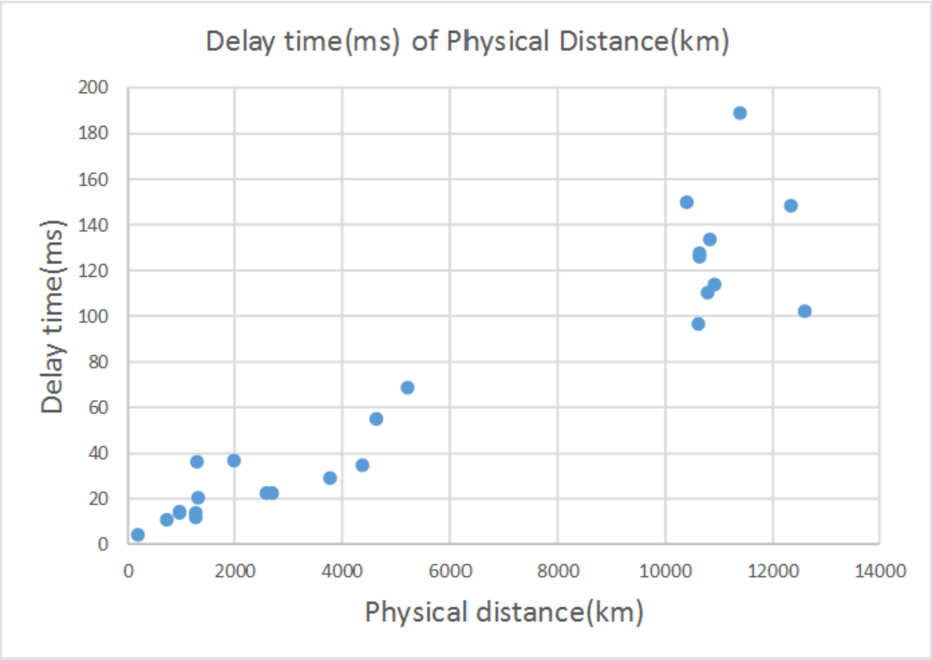
  -m: the max time-to-live used in outgoing probe packets;
  -n: print hop addresses numerically rather than symbolically and numerically;
  -p: the base UDP port number used in probes;
  -q: TTL test number;
  -r: Bypass the normal routing tables and send directly to a host on an attached network;
  -s: Use the following IP address as the source address in outgoing probe packets;
  -t: set the type-of-service in probe packets to the following value, the value must be a decimal integer in the range of 0 to 255;
  -f: Set the initial time-to-live used in the first outgoing probe packet;
  -v: Verbose output. Received ICMP packets other than time_exceeded and unreachables are listed;
  -w: Set the time to wait for a response to a probe.
- Example: tracert -d -h maximum_hops -j host-list -w timeout target_host:
  - -d: spcecifies to not resolve addresses to host names;
  - -h mamimum_hops: maximum number of hops to search for the target;
  - -j host-list: specifies loose source route along the host-list;
  - -w timeout: waits the number of milliseconds specified by timeout for each reply;
  - Target_host: specifies the name or IP address of the target host.

(b) Traceroute study:

I implemented some websites and recorded their delay time, hops and physical distance between the servers. Here is the table:

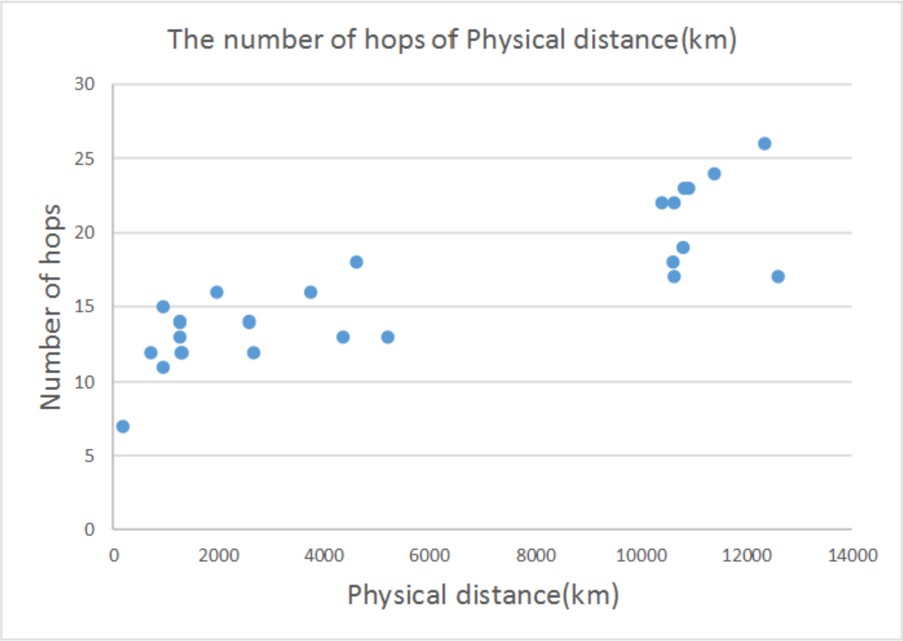
Number	Website	Delay time(ms)	Hops	Physical Distance(km)
1	www.upei.ca	4.500166667	7	187.5
2	web.mit.edu	10.62816667	12	713.7
3	www.nyu.edu	14.41333333	11	950.92
4	www.google.com	14.05066667	15	960.57
5	www.cbc.ca	11.65833333	13	1265.9
6	www.utoronto.ca	13.98616667	14	1265.9
7	www.washington.edu	36.13766667	12	1282.2
8	www.wikipedia.org	20.3065	12	1307.4
9	www.apple.ca	36.8995	16	1967.9
10	www.umanitoba.ca	22.59766667	14	2579
11	twitter.com	22.328	12	2673.7
12	www.ucalgary.ca	29.32666667	16	3760.5
13	www.ubc.ca	34.61766667	13	4357.3
14	www.cam.ac.uk	54.98166667	18	4619.1
15	www.ub.edu	68.89916667	13	5215.7
16	www.neu.edu.cn	149.7308333	22	10389.3
17	www.sina.com	96.54625	18	10596.7
18	www.tsinghua.edu.cn	127.4403333	17	10622.2
19	www.pku.edu.cn	126.0026667	22	10629.6
20	www.u-tokyo.ac.jp	110.1058333	19	10789.5
21	www.uos.ac.kr	133.657	23	10821.4
22	www.sohu.com	114.0516667	23	10899.3
23	blog.csdn.net	188.68025	24	11385.8
24	www.xmu.edu.cn	148.3038333	26	12343.6
25	www.baidu.com	102.1968333	17	12585.7

The plot of delay time(ms) of physical distance(km)



According to the plot, we can find that when delay time is increase, the physical distance also increases. Therefore, we can conclude that the delay time has positive correlation with physical distance.

The plot of the number of hops of physical distance(km)



According to the plot, we can find no matter how far the physical distance it is, the number of hop does not have big change. Therefore, we can conclude that the physical distance does not have strong influence to number of hops.

Conclusion: the physical distance has significant influence to the delay time, the more physical distance it is, the much delay time it has. Besides, the number of hops does not be influenced by physical distance.

2.

- Title
  Extracting Benefit from Harm: Using Malware Pollution to Analyze the Impact of Political and Geophysical Events on the Internet
- Abstract
  People use Internet background radiation(IBR) to research the internet malicious activity for decades' year. In this paper, the author gives us an analyze of using the same method to evaluate normal internet activities. Two phenomena the paper is focus on: one is restriction of Internet communications based on the country level; another one is concentrate on natural disasters. A brand-new local IBR activity metric is introduced. The principle of metric is calculating the number of "unique IP addresses per hour contributing to IBR". [1]
- Introduction
  IBR is a voluntary one-way network which is sent to random IP addresses. That addresses have software behaviors with unintended or malicious. There is a giant background radiation which is generated by computer, however, mostly the computer has not information of their users. Besides, it is not obviously enough to obtain the characteristics of the malware by its data. In this paper, the author illustrates that they analyze IBR traffic to yield judgement of affection of internet connectivity but not only related to malware. According to Pang's paper [2], three primaries can cause IBR traffic: backscatter attacks, scans and bugs. Backscatter attacks means backscatter from spoofed denial-of-service(DoS) attacks. More specifically, DoS attacks try to crush the user with traffic or agreement so that they can reduce users' ability of legal requests. System misconfiguration of IBR traffic means user set a wrong IP address. The platform can both be DNS or some server. The IBR traffic can be detected by network telescopes. Which is usually happen is the unknown server IP addresses, so it is also called darknets. In commerce area, people use darknets to keep data safe.
  In this paper, the author obtained the data from UCSD network telescope because the UCSD's captures can reach 2-10GB approximately per hour. In the future research, those data can also be used or sharing.
  First, they make an overview of analyzing macroscopic actions by internet data sources. Then they give a conclusion of their recent research: the censorship experiences around two countries' internet. Next, an introduce will be gave about earthquake analyzing by IBR traffic, one is New Zealand in February 2011, the other one is in March same year of Japan. They share some points, including strongly powerful, which damaged necessary infrastructure, and destroyed millions of computers which are accessible to internet and cut down the IBR traffic. In the end of paper, the limitation of the approach will give and discussion of future research will have.
- Related work and paper
  The related work is doing some research of relationship of earthquake and internet macroscopic effects.

The first thing they discussed is border gateway protocol(BGP). In BGP internet routing, the information is announced and withdraw by Autonomous Systems(ASes). They are departed from prefix of network and use to the rest of network. Sometimes the number of prefixes may change because of the world wide propagation, which is unnecessary to find those changes. Using BGP to detect physically observations, they are focus on the different BGP attributes per minute bins BGP message's exchange, and detect the difference between profile definition and reference.

Active probing is a tool of solving IP layers problem. Sometimes it is used to help analyze BGP. [3] The meaning of analyses based is entirely focus on active probing data. Probing data can be used on learning global internet topology. [4,5,6,7]

Passive traffic measurement is use for analyzing macroscopic internet action. However, most of traffic data are unavailable to researchers to capture.

Google service may be more accurately, which make more interesting data is available. It is the main data source which researchers used. Google like a library of human behavior on internet, which can point out the network state indirectly.

Peer to peer traffic is another data source. Peer to peer(P2P) is a popular definition in recent year. It also including numerous data sources. in the past, IBR traffic is always generated by malware because of users' end actions, therefore, the researcher abandons this.

Internet background radiation is proposed by Casado et al. [8] they called IBR as "opportunistic measurement" because they thought IBR makes internet measurement data viable.

- Censorship

  According to BGPmon website [9], the Egyptian networks are stopped on January 27th, 2011 which lasted five days. The BGP IPv4 is withdraw from Egypt, this event even led to President resigned. Researchers analyzed three kinds of data and found the darknet which mentioned above has significant difference based on the vary behaviors. If the traffic rate wants to be isolated influence of host count, the prior option is exam IBR traffic subset. Review the event, it can be found that some IBR traffic still connect during the disaster, there are two reasons may enable to explain: one is some network prefixes were still active; the other one is the outage did not affect inbound connection, which means the outbound connection can still work. Also, it can be found not only unique IP addresses, but also packets' rate was decrease gradually because the continuous of withdraw of more routes.

- Earthquakes

  Looking the earthquake of New Zealand and Japan, those cases represented the influence of IBR in natural disaster. A metric is helpful to analyze the impact. Based on the contrast, it can be found the distance between the earthquake center of Tohoku and some important IP addresses; and the relationship between distance and population. The histogram graph is also used to analyze the different distance between epicenter of earthquake and network address ranges. Then the earthquake impact maximum radius can be estimated by network connectivity.

  A similar analyzation took on New Zealand's earthquake [10] as well. After those research, the scientists found the peak of IP addresses per hour are always has 140 on weekends, while the maximum number may reach 160. After the earthquake, within 24 hours, the rate will decrease around 100 IPs per hour. Then the rate will increase slowly and it will take a week and return the previous level.

- Results

  According to geolocating the traffic destination of source IP addresses to the darknet addressed, the lost connectivity can be found in specific region. On the other hand, there are some limitation exist. First, some unpredictable things may influence the IBR's data source reliable. Second, sending traffic to darknet may have inaccuracy elements. In the future this problem may have more improvement.

- Reference:

[1]Dainotti, A., Amman, R., Aben, E., & Claffy, K. C. (2012). Extracting benefit from harm: using malware pollution to analyze the impact of political and geophysical events on the Internet. *ACM SIGCOMM Computer Communication Review*, *42*(1), 31-39.
[2] Pang, R., Yegneswaran, V., Barford, P., Paxson, V., & Peterson, L. (2004, October). Characteristics of internet background radiation. In Proceedings of the 4th ACM SIGCOMM conference on Internet measurement (pp. 27-40). ACM.
[3] J. Cowie. Egypt leaves the internet. http://www.renesys.com/blog/2011/01/egyptleaves-the-internet.shtml, January 27 2011.
[4] CAIDA. Archipelago measurement infrastructure. http://www.caida.org/projects/ark/, July 28 2011.
[5] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani. iPlane: An Information Plane for Distributed Services. OSDI, 2006.
[6] RIPE NCC. RIPE Atlas. http://atlas.ripe.net/.
[7] Y. Shavitt and E. Shir. DIMES: Let the Internet Measure Itself. ACM SIGCOMM Computer Communication Review, October 2005.
[8] M. Casado, T. Garfinkel, W. Cui, V. Paxson, and S. Savage. Opportunistic Measurement: Spurious Network Events as a Light in the Darkness. In ACM Fourth Workshop on Hot Topics in Networks (HotNets-IV), New York, NY, USA, 2005. ACM.
[9] BGPmon. Internet in Egypt offline. http://bgpmon.net/blog/?p=450, January 28 2011.
[10] N. Mathewson. Power restored to most households. http://www.stuff.co.nz/national/christchurchearthquake/4734825/, March 5 2011.

3. propagation delay: $d_1 = \frac{m}{s}$

Transmission delay: $d_2 = \frac{L}{R}$

$d_1 = d_2$, therefore, $\frac{m}{s} = \frac{L}{R}$, $\frac{m}{2.5\times10^8} = \frac{100}{28}$, then we can get $m \approx 892{,}857{,}143\text{m} \approx 892.857\text{km}$.

4. (a) On virtual- circuit packet-switched, let's assume the time of transmission is t_1,

$$t_1 = \frac{h+F}{R} \times Q.$$

Because the VC packet-switched has set up time t_s, therefore, the total time of sending packet from source to destination is:

$$t = t_1 + t_s = \frac{h+F}{R} \times Q + t_s$$

(b) the datagram packet-switched network is connectionless service, therefore,

$$t = t_2 = \frac{2h + F}{R} \times Q$$

5.(a) for one big message:

$$t = t_{prop} + t_{trans} = \frac{m}{s} + \frac{L}{R} = \frac{20000 \times 10^3 m}{2.5 \times 10^8 meters/sec} + \frac{80000bits}{2000000bps}$$

$$= 0.08s + 0.4s = 0.48s$$

(b) For acknowledge packet:

$$t_{ack} = t_{prop} + t'_{trans} = 0.08s + 0.1s = 0.18s$$

For packet:

$$t_{packet} = t_{prop} + t''_{trans} = 0.08s + \frac{L'}{R'} = 0.08s + \frac{40000bits}{2000000bps} = 0.1s$$

For 20 packets transmission:

$$t' = (t_{ack} + t_{packet}) \times 20 = (0.1s + 0.18s) \times 20 = 5.6s$$

6.(a) $t_{total} = t_{prop} + t_{handshaking} + t_{tran} = \frac{RTT}{2} + 2 \times RTT + \frac{1.5MB}{10Mbps}$

$$= \frac{0.08s}{2} + 2 \times 0.08s + \frac{12582912bits}{10000000bps} \approx 1.458s$$

(b) First, we count the number of packets needed to send the file:

$$\frac{1.5MB}{1kb} = 1536bits;$$

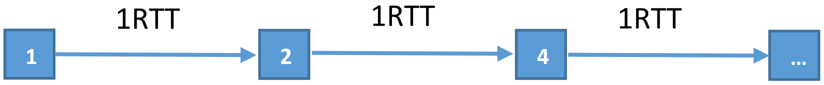
Therefore, we need (1536-1) x RTTs to be added to the total delay time.

$$t'_{total} = t_{handshaking} + t_{tran} + \#packets \times t_{prop} = 2 \times RTT + \frac{RTT}{2} + \frac{1.5MB}{10Mbps} + 1535bits \times RTT = 124.26s$$

(c) divide the 1536 packets to 20 packets, $1536 \div 20 = 76.8$, and the propagation time is ignored. Therefore,

$$t''_{total} = t_{handshaking} + t_{tran} = 2RRT + RRT \times (76.8 - 1) = 77RRT = 6.28s$$

(d) the process of sending packet is:



Therefore, there are $1 + 2 + 4 + \cdots + 2^n = 2^n - 1$ packets. When n=11, can we send all packets.

$$t'''_{total} = t_{handshaking} + t_{tran} = 2RRT + RRT \times 11 = 13RRT = 1.04s$$

7. FTP connection: choose 50000 as source port and 21 as destination port;

SSH connection: choose 51000 as source port and 22 as destination port;

(a) s, r11, S, D, 50000, 21;
(b) r14, r21, S, D, 50000, 21;
(c) r22, r34, S, D, 50000, 21;
(d) r31, d, S, D, 50000, 21;
(e) r11, s, D, S, 21, 50000;
(f) r21, r14, D, S, 21, 50000;
(g) r34, r22, D, S, 21, 50000;
(h) r14, r21, S, D, 22, 51000;
(i) r22, r34, S, D, 22, 51000;
(j) r31, d, S, D, 22, 51000;
(k) r11, s, D, S, 51000, 22;
(l) r21, r14, D, S, 51000,22.