CSEE 4119: Computer Networks — Due on Thursday, February 6th, 2014

Assignment #1 – Warming up with Computer Networks

P. Prakash, W. An, P. Nirantar, R. Yu, X. Zhu (TAs) A. Chaintreau (instructor)

How to read this assignment **:** Exercise levels are indicated as follows

- ($\rightarrowtail$) "elementary": the answer is not strictly speaking obvious, but it fits in a single sentence, and it is an immediate application of results covered in the lectures.

  Use them as a checkpoint: it is strongly advised to go back to your notes if the answer to one of these questions does not come to you in a few minutes.

- ($\curvearrowright$) "intermediary": The answer to this question is not an immediate translation of results covered in class, it can be deduced from them with a reasonable effort.

  Use them as a practice: how far are you from the answer? Do you still feel uncomfortable with some of the notions? which part could you complete quickly?

- ($\looparrowright$) "tortuous": this question either requires an advanced notion, a proof that is long or inventive, or it is still open.

  Use them as an inspiration: can you answer any of them? does it bring you to another problem that you can answer or study further? It is recommended to work on this question only AFTER you are done with the rest!

Exercise 1: First programming warm up (15 pt)

Motivation: Get to know the Wireshark tool.

1. ($\curvearrowright$) Complete the first Wireshark Lab (see additional document provided with the assignment).

Exercise 2: Statistical Multiplexing (8 pt)

Motivation: In class, we saw that packet switching can make more efficient use of resources by taking advantage of the fact that only a fraction of potential senders are active at any time. In this problem, you will be asked to demonstrate this fact mathematically.

Suppose we have a single link with capacity L bits per second and a population of users that generate data at r bits per second when busy. The probability that a user is busy generating data is p

1. ($\rightarrowtail$) What is the maximum number of users that can be supported using circuit switching? Call this value MC.

2. ($\curvearrowright$) Now suppose we use packet switching to support a population of MP users. Derive a formula (in terms of p, MP, N, L, and r) for the probability that more than N users are busy? (Hint: Start by deriving a formula for the probability that exactly N users are busy.)

3. ($\curvearrowright$) Plug in some numbers. Let L = 1Gbps, r = 64Mbps and p = 0.2. Give the value for MC. What is the probability that more than MC users are busy for MP = 2 $\times$ MC? What about MP = 4 $\times$ MC? [Note: Feel free to write a quick program, or use Mathematica or Excel to compute these numerical values.]

Exercise 3: Assuming you were a packet, how fast would you make it here? (4 pt)

1. $(\curvearrowright)$ What is the propagation and transmission delay on a 1 Gb/s fiber optic circuit from Ulaanbaatar to New York? [Hint: try http://www.indo.com/cgi-bin/dist].

   To answer this question, we assume that the circuit is made of eleven nodes, including the sender in Ulaanbaatar and the receiver in New York, and that it follows the shortest path between the two points (on the surface of the earth). We assume that the packet length is 1,500 bytes and that the propagation speed on the fiber is $0.66 \times c$.

   Can you then estimate the total delay for a packet to cross this circuit?

2. $(\curvearrowright)$ For which range of packet size would it be faster to take this information with you on a plane that travels at 885 km/h? What if you include in addition, a four hours delay to pass security, customs and take the "A" line to the campus? What if you travel by horse, which would take approximately 6 months[1]? Can you estimate the weight of a hard drive, DVD or flash that you need and how many horses in addition to yours are needed to carry those?

Exercise 4: More on delays (8 pt) This elementary problem begins to explore propagation delay and transmission delay, two central concepts in data networking. Consider two hosts, Hosts A and B , connected by a single link of rate R bps. Suppose that the two hosts are separated by m meters, and suppose the propagation speed along the link is s meters/sec. Host A is to send a packet of size L bits to Host B. No transmission error occurs.

1. $(\rightarrowtail)$ Express the propagation delay, d_{prop} in terms of m and s.

2. $(\rightarrowtail)$ Determine the transmission time of the packet, d_{trans} in terms of L and R .

3. $(\rightarrowtail)$ Ignoring processing and queuing delays, obtain an expression for the end-to-end delay.

4. $(\curvearrowright)$ Suppose Host A begins to transmit the packet at time t = 0 . At time t = d_{trans}, where is the last bit of the packet?

5. $(\curvearrowright)$ Suppose d_{prop} is less than d_{trans}. At time t = d_{trans}, where is the first bit of the packet?

6. $(\curvearrowright)$ Suppose s = 4.5×10^4, L = 1,000 bits and R = 22Mbps. Find the distance m so that $d_{\text{prop}} = d_{\text{trans}}$.

[1] 1 month to go to Beijing, two weeks to traverse the Pacific and 4.5 months to go from San Francisco to New York.

Exercise 5: The impact of layering on innovation (4 pt)

Motivation: A very important debate in the networking research community today is whether the protocol stack that we have adopted is an impediment to innovation in computer networking.

1. ($\rightarrowtail$) Which one of the following devices need to implement functions from the Link Layer (*i.e.*, read the link layer header of a frame to do some operations) ? which ones for the Transport layer? which ones for the Application layer?

   {a router in a tier-1 ISP, an ipad, a link-layer switch, a router in a tier-3 ISP, a laptop, a web-server}

2. ($\curvearrowright$) Several versions of BitTorrent have been written and have been quickly adopted. The TCP protocol went once through a series of revision in a couple of years. The transition from IPv4 to IPv6 (a new version of a network layer protocol) is still not made more than a decade after its standardization.

   Using elements from the lectures and inspired by the previous question, can you explain why some new protocols get adopted quickly while others take a long time?

Exercise 6: When measuring the delay may confuse you (1 pt)

Motivation: Understand what can go wrong when results from traceroute are taken too literally

Traceroute works by sending "probes", which is a packet with a limited "Time-To-Live" (TTL) value indicated inside the networking header. Remember that the networking header is the part that the router reads and interpret when it receives a datagram. Hence, a router on a path to a destination, once it realizes that the packet already traversed exactly TTL links, do not forward the packet any further but instead it sends an answer to the source with its hostname, IP address, and a unique packet signature. Traceroute first sends a packet with TTL indicated as 1, it shows the result (hostname, IP address, and the delay to receive this answer. It then sends a new probe with TTL indicated as 2, and so on. This tool was initially created for debugging purpose. It allows to estimate:

- which sequence of routers, and hence which links, are used on a path used to connect to a destination.
- measure the delay between the source and each intermediate hop.

In this exercice, to simplify we assume that traceroute only sends a single probe packet.

1. ($\looparrowright$) Describe a scenario where results from traceroute may incorrectly induce you to conclude that two routers are connected by a direct link? or even that a router may have looping link (link from the router to itself)?

2. ($\looparrowright$) Can you comment on partial or complete solutions to eliminating these problems? As an example, do you think it's possible to build a new version of traceroute that allows to learn correctly the real links that could be used between two hosts?