Networked Systems, COMPGZ01, 2009

Answer TWO questions from Part ONE on the answer booklet containing lined writing paper, and answer ALL questions in Part TWO on the multiple-choice question answer sheet.

Marks for each part of each question are indicated in square brackets

Calculators are permitted

Part ONE

1. **Intra-Domain Routing (and a bit of TCP)**

   Consider a large-scale single routing domain that uses distance-vector routing as its intra-domain routing protocol. In this domain, all link weights are set to one. Assume that the routing protocol implementation restricts link weights to be positive integers.

   a. Suppose you gain control of one router in the domain, and can modify the software of that one router in any way you like. Suppose you want to "hijack" the traffic destined to one particular subnet in the routing domain; that is, you would like to cause traffic destined to this victim subnet to be routed to the router you control, rather than the correct router (*i.e.,* the router actually connected to the victim subnet).

   Describe a simple attack you can make, whereby altering the distance-vector announcements of the router you control, you can force some other routers to forward traffic destined for the victim subnet to the router you control.

   [8 marks]

   b. When you carry out the attack you just described, *which* routers in the routing domain will be forced to send victim-destined traffic to "your" router? Be specific, based on the way distance-vector routing works.

   [7 marks]

c. The manager of this campus network would like to improve capacity by allowing routers to forward traffic over *multiple paths*. That is, rather than forwarding all traffic for a destination to the same next hop, a router may choose to spread the traffic over several *link-disjoint* paths, and thus be able to use greater total traffic capacity. (Two paths are link-disjoint if they share no links in common.)

   The network manager says that distance-vector routing won't support multi-path routing well, and that she must instead change to link-state routing. Assume that the routing protocols themselves cannot be changed from the standard definitions of these two protocols' behavior. That is, the *forwarding* behavior at the routers can be modified to support multi-path routing, but it must only make use of knowledge gained by participating in the standard distance-vector or link-state protocol.

   Is the manager right? Does distance-vector routing itself allow multi-path routing on link-disjoint paths? Does link-state routing? For each, explain why or why not.

   [8 marks]

d. After deploying multi-path routing on the campus network, the network manager notices that TCP performance plummets between one particular source-destination pair connected by multi-path routes. The two link-disjoint paths used are 3 and 10 hops long. She determines there is no congestion on either of these paths, and that all the packets transmitted by the sender do arrive at the receiver.

   Why has TCP performance on this multi-path route decreased? Refer specifically to a mechanism that is part of TCP in your answer.

   [10 marks]

[Total 33 marks]

2. **Synchronizing in White Space**

Having just heard in lecture about the wireless spectrum freed up by the transition to digital television, GZ01 student Radia Hertz is designing a digital packet radio system for the resulting radio "white spaces."

Hertz's first task is to program her white space receiver to synchronize its bit clock with an incoming packet transmission. For this, she decides to use the Manchester encoding algorithm presented in lecture. Suppose her transmitter takes a stream of data bits, and directly applies the Manchester encoding algorithm to produce an output stream of Manchester-encoded bits.

a. How would Hertz's transmitter encode a packet consisting of five 1s? (Your answer should be a string of bits).

[2 marks]

b. How would Hertz's transmitter encode a packet consisting of five 0s? (Your answer should be a string of bits).

[2 marks]

At the receiver, assume that Hertz's physical layer (PHY) design has a method of detecting the beginning and end of a packet, with an uncertainty of plus or minus one Manchester-encoded bit. Furthermore, if the PHY indicates start of packet one bit early (or end of packet one bit late), the extra bit it passes up is, non-deterministically, a zero with probability 1/2 and a one with probability 1/2.

Thus, for example, if P is a Manchester-encoded packet and the receiver's PHY indicates start of packet one bit early and end of packet one bit late, the receiver will pass up one of four possibilities with equal probability: 1P1, 0P1, 1P0, or 0P0. If the receiver's PHY indicates start of packet one bit early and at the correct end of packet time, the receiver will pass up one of two possibilities with equal probability: 1P or 0P.

After the receiver's PHY passes up the Manchester-encoded packet, the receiver's link layer searches through the packet for valid Manchester-encoded bits, decodes the entire packet, and passes it up to the network layer.

c. Under some circumstances a problem can result from the interaction between Hertz's PHY design and her Manchester encoding, where the network layer at the transmitter sends one bit string, but the network layer at the receiver passes up a **different** bit string. Give an example of this behavior by specifying the following:

   i. A bit string (of your choosing) at the sender's link layer.
   [2 marks]

   ii. The resulting bit string at the sender's PHY.
   [2 marks]

   iii. A possible resulting bit string passed up by the receiver's PHY.
   [2 marks]

   iv. The corresponding bit string passed up by the receiver's link layer.
   [2 marks]

d. Are all payloads vulnerable to the data corruption bug, or only some? If all payloads are vulnerable, explain why; if only some are, explain which, and why those payloads are vulnerable to data corruption.

[5 marks]

e. Without modifying the PHY, briefly describe a link layer fix that will guarantee correct data delivery.

[3 marks]

Radia follows your advice, and has a mostly-working link layer/PHY sending frames in the white space spectrum. Every now and then, however, interference from other transmissions causes corrupted packets to be passed up the stack. To lessen the liklihood of this occurring, Radia decides to add a three-bit CRC check to each frame.

She uses the CRC algorithm as described in lecture, with generator bits G = 1001.

f. Given a generator G and data bits D, describe in words how the CRC algorithm calculates the CRC bits R.

[3 marks]

g. Given data bits D = 101110000, compute the CRC bits R that will be appended to the end of the frame.

[5 marks]

h. What is the false positive liklihood of Radia's CRC check? In other words, suppose when the receiver computes the CRC of the received data and compares it to the CRC value that the sender appended to the frame, what is the likelihood that the two are equal, yet the data has been corrupted over the air? (Assume random data payloads.)

[5 marks]

[Total 33 marks]

3. **TCP (and a bit of 802.11 and the End-to-End Argument)**

a. Suppose that two TCP flows share the same congested link, but follow different end-to-end paths, such that one flow has a round-trip time (RTT) of 40 ms, and the other has an RTT of 200 ms. Assume that the congested link offers a *constant* packet drop rate; it does not change as either of these two TCP senders varies its sending rate.

One of these two TCP flows will achieve higher throughput than the other. Which one, and why? Refer to the relevant details of TCP's congestion control algorithm in your answer.

[6 marks]

b. Suppose a TCP sender sends across a network path that is otherwise unused; this one sender is the sole user of all the links along the path. The links along the path may vary in their bandwidths, however. Ignoring slow start and congestion control, what *fixed* send window size would the TCP sender need to use in order to achieve the full bandwidth allowed by the path? Explain your answer from first principles using the relevant details of TCP sender behavior.

[6 marks]

c. Describe *slow start* as implemented by the TCP sender. In your description, include:

- the problem that slow start solves
- the events that cause the sender to trigger slow start
- how the sender's window size changes during slow start
- when the sender exits slow start

[7 marks]

d. Using a Chiu-Jain plot, illustrate the efficiency and fairness behavior of multiplicative-increase, multiplicative-decrease (MIMD) congestion control. Use the diagram to explain why MIMD congestion control does not converge to a desirable combination of efficiency and fairness.

[7 marks]

e. The Internet's transmission control protocol (TCP) provides reliable delivery between Internet-attached end hosts. The 802.11 wireless medium access control (MAC) layer also implements reliable delivery at the link layer.

   Given that most traffic users send over 802.11 links is TCP traffic, is it appropriate for *both* TCP and the 802.11 link layer to implement reliability, and why or why not? Apply the end-to-end argument in your answer.

[7 marks]

[Total 33 marks]

4. **Huffman coding from Bloomsbury to Boston**

Well-known London writer and GZ01 student Virginia Woolf is interested in transmitting her novels via Morse code across the Atlantic in the most efficient manner possible. As her GZ01 classmate, she enlists your help in this endeavor. Virginia points out that the letters in English text don't occur equally often—some occur much more frequently than others. Consequently, she suggests Huffman coding. To begin the project, you both decide to refresh your understanding of Huffman coding by constructing the Huffman code for just the vowels, using their relative frequencies of occurrence in English text.

Recall that the algorithm for Huffman coding presented in lecture maintains two data structures: a directed binary tree graph G whose edges are labelled with either "1" or "0," and a priority queue S whose entries are pairs of node names and relative frequencies.

You look up the relative frequencies of occurence in English text for vowels, and find the following:

Vowel	Frequency (Pct.)
A	21
E	33
I	19
O	20
U	7

a. How much information (measured in bits) per vowel is there in a string of vowels chosen to occur with the frequency distribution shown above?

[5 marks]

b. How does the Huffman encoding algorithm initialize S and G, given the above as input?

[2 marks]

c. The Huffman algorithm then makes four iterations as described in lecture. In your answer booklet, write the contents of S and G at the end of each of the four iterations.

[10 marks]

d. How many bits per vowel does your Huffman code require, for a randomly-chosen string of vowels that occur with the same frequency as they do in English text?

[4 marks]

e. In general, what is the most important reason that Huffman coding is suboptimal for the problem of encoding English text?

[4 marks]

Pleased with your work, Virginia looks up the relative frequencies for the consonants (i.e., those letters that are not vowels), and constructs a separate Huffman tree for them. She mails the Huffman trees across the Atlantic to Boston, and then using the binary edge labels on the two Huffman trees, begins tapping out her novels in Morse code over the telegraph. You burst into the room, telling her to stop.

f. What is the problem here—why might Boston have a hard time decoding Virginia's messages?

[4 marks]

g. One fix to this problem is to reencode the entire alphabet in one big Huffman tree. Is this more, less, or equally efficient than the scheme that you and Virginia have designed? Explain why.

[4 marks]

[Total 33 marks]

Part TWO

Multiple Choice

The questions in this section are multiple choice. Zero, one, or more than one choice may be correct, and the number of correct choices varies from question to question. Using an HB pencil, fill in the rectangle on the multiple choice answer sheet corresponding to the letter(s) of ALL correct choices.

Do not guess. You will be awarded one mark for each choice correctly filled, but you will lose one mark for each choice incorrectly filled. Final results will be normalized across the class.

1. **Layering**

   Which of the following are true statements about a layered protocol design? Assume in the following that we number higher layers with higher numbers.

   A. Communication takes place only between adjacent layers.
   B. Layer $n+1$ knows nothing about the content of layer n's payload.
   C. Layer n encapsulates layer $n+1$'s header and payload with its own header.
   D. Protocol headers allow demultiplexing to higher layers.

2. **Subnetting and supernetting**

   Circle all true statements:

   A. Subnetting allocates multiple contiguous network numbers to one physical network.
   B. CIDR allows multiple physical networks to share a single network number.
   C. IP addresses allocated to a CIDR network don't have the two layer hierarchial structure (network number, host number) of non-CIDR IP addresses.
   D. NAT boxes need to keep CIDR prefix length state for hosts behind the NAT with CIDR-assigned IP addresses.
   E. The wide deployment of CIDR allows routers on the Internet to use shortest-prefix matching when they use a datagram's IP address to lookup the output port to send it to.

3. **The Akamai content distribution network**

Recall the following parts of the Akamai Content Distribution Network (CDN) architecture described in lecture:

- TLNS: Top-level nameserver
- LLNS: Low-level nameserver

Circle all true statements:

A. Since DNS A records for TLNSs have TTLs of 30-60 minutes, upon LLNS failure, clients fail over to a different LLNS in at most 60 minutes.

B. Since DNS A records for LLNSs have TTLs of two minutes, upon a single edge server failure, clients fail over to a different edge server in exactly two minutes.

C. Akamai's CDN masks origin web server (content server) failures by replicating the TLNS in many different networks worldwide.

D. To achieve resiliency to network failure, TLNSs return LLNSs in different regions (ASes).

E. To measure latency between different regions, Akamai servers establish peering sessions with BGP routers.

4. **Error detection and correction**

Consider the following scheme for detecting and correcting bit errors, called *3-repetition coding*. The sender takes input bits one at a time, and outputs three copies of each input bit. Thus, k input bits $b_0 b_1 \ldots b_{k-1}$ result in the message

$$m = b_0 b_0 b_0 b_1 b_1 b_1 \ldots b_{k-1} b_{k-1} b_{k-1},$$

of total length $n = 3k$.

Circle all true statements:

A. By voting, 3-repetition coding can correct any and every set of bit errors to m such that no pair of adjacent bits is errored.

B. 3-repetition coding can detect any single-bit error in m.

C. $d_{\min}$ (the minimum Hamming distance between any pair of codewords) for the 3-repetition code is 3.

D. The 3-repetition code is a block code.

E. The rate of the 3-repetition code is 3.

5. **Ethernet**

Circle all true statements pertaining to Ethernet as described in lecture and assigned reading "Ethernet: Distributed Packet Switching for Local Computer Networks" by Metcalfe and Boggs.

Let τ represent the propagation delay from one end of an Ethernet to the other. Assume that the Ethernet and stations are not damaged or otherwise malfunctioning.

A. Transmissions beginning more than τ apart are guaranteed not to result in collision.

B. Once an Ethernet station has been transmitting for 2τ, it is guaranteed not to hear that its transmission has collided.

C. Once a station has synchronized to another's transmission with Manchester encoding, its backoff slots are time-aligned with the other station's to within $\tau/2$ or less.

D. Whether a jamming sequence is needed to enforce consensus on collisions depends only on bit rate and τ.

6. **Reliable data transfer**

Circle all true statements:

A. When guaranteeing exactly-once delivery, window-based reliable transfer protocols may increase window size to size of the sequence number space minus one.

B. A selective-repeat (SR) receiver sends cumulative ACKs back to the sender when its receive window increases in size.

C. The stop-and-wait protocol is identical to the Go-Back-N protocol with a window size of one.

D. The selective-repeat protocol can guarantee exactly-once delivery in finite time.

7. **HTTP-TCP interaction**

Which of the following could improve the performance of a web browser that opens a new TCP connection for each HTTP transaction?

A. Enabling Nagle's algorithm on the TCP connections the browser opens.

B. Using persistent HTTP connections to amortize the burden of TCP slow start.

C. Raising the initial TCP retransmission timeout (`rto`).

D. At the web server, transforming multiple `write` system calls each with small amounts of data into one `write` system call with a larger amount of data.

8. **BGP Routing**

Which of the following statements about the Border Gateway Protocol (BGP) are correct?

A. BGP may choose routes that contain loops immediately after a topology change.

B. Given three autonomous systems (ASes), A, B, and C, consider the round-trip time (RTT) of the path selected by BGP between A and B with the sum of the round-trip times of the paths selected by BGP between A and C and between C and B. BGP never selects paths such that the RTT of the former path ($A \rightarrow B$) is greater than the sum of the RTTs of the latter paths ($A \rightarrow B$ and $B \rightarrow C$).

C. If one tier-1 ISP stops peering with another tier-1 ISP, no customers of either ISP will become partitioned from one another, because BGP will then begin to route traffic between these two tier-1 ISPs via some third tier-1 ISP.

D. BGP routers in an AS must honor multi-exit discriminator (MED) attributes advertised by routers in neighboring ASes.

E. Under BGP, routers not on an AS's border (*i.e.,* those connected only to other routers within the same AS, with no links to routers in other ASes) need not store routing tables with entries for all destination prefixes in the global Internet.

9. **802.11 MAC**

Consider the IEEE 802.11 medium access control (MAC) protocol. Which of the following statements are true?

A. Request to Send (RTS) and Clear to Send (CTS) packets are of no use in mitigating the hidden terminal problem when senders attempt to send very short data packets.

B. 802.11 senders use collision detection to share the wireless medium more efficiently.

C. Suppose two 802.11 senders, A and C, wish to send concurrently to two 802.11 receivers, B and D, respectively. Carrier sense may *reduce* the aggregate throughput achieved by A and C in some topologies, as compared with the aggregate throughput achieved when neither uses carrier sense.

D. Address Resolution Protocol (ARP) request packets sent on 802.11 networks are always acknowledged at the link layer.

E. An 802.11 sender backs off exponentially before retransmitting when it does not receive a link-layer acknowledgement (ACK) for its unicast data transmission.

10. **Network Security**

A. Slammer spread so quickly because it was a topological worm, rather than a random-scanning worm.

B. A firewall that disallows inbound TCP connections also prevents conventional *outbound* ftp sessions.

C. When a worm first begins spreading using random scanning, the population of infected hosts initially grows linearly over time.

D. A network telescope sends probe traffic to remote networks to measure how many hosts have been infected by a worm.

E. If a firewall is able to parse UDP headers, but not headers at protocol layers above UDP, it cannot use stateless filtering to block inbound UDP connections.

11. **TCP Protocol and Congestion Control**

A. TCP practices self-clocking transmission by using a timer at the sender to pace packet transmissions at a perfectly smooth rate.

B. Over a network that can drop packets, the TCP protocol ensures using FIN messages that once both endpoints of a TCP connection decide to end the connection, they will always both reach agreement that both have reached this decision.

C. For very short transfers of fewer than 4 packets, TCP cannot invoke fast retransmit to recover from losses.

D. In normal operation, a TCP sender repeatedly provokes loss of its packets at one or more queues along the path between sender and receiver.

E. TCP begins new connections with an initial sequence number of zero.

[Total 33 marks, after normalization of marks]

END OF PAPER