CNT5106C Computer Networks, Fall 2014
Instructor: Prof. Ahmed Helmy
Homework #3
On the Data link layer, MAC protocols and Wireless Networking
[Date Assigned: Nov 19th, 2014. Due Date: Tues Dec 2nd, 11:55pm]
Total Points: 132, plus 20 extra points

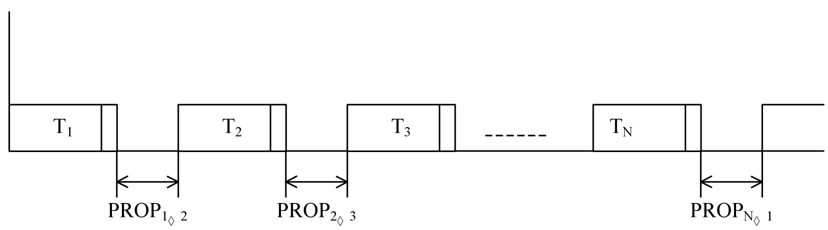
Q1. (17 points: 9 + 3 + 5)

I. Derive the utilization formula for: (a) Ethernet, (b) Token ring (release after reception), and (c) FDDI [which uses Token ring (release after transmission)].

II. Compare them.

III. Comment on what happens when the number of stations on the LAN is increased in each case.

A1. I. For Ethernet and token ring release after reception (RAR) the derivation is given in the lecture. (2.5 points for each)
For token ring release after transmission (RAT): (4 points)



Utilization

$$\mu = \frac{Useful\,Transmission\ \ time}{Total\ \ time(Useful + wasted)} = \frac{T_1 + T_2 + \ldots\ldots + T_N}{T_1 + T_2 + \ldots\ldots + T_N + \sum PROP_{i \to i+1}} = \frac{\sum T_i}{\sum T_i + PROP}$$

$$a = \frac{PROP}{E(T_n)}$$

where $E(T_n)$ = expected (average) duration of node transmission $= \sum_i^N \frac{T_i}{N}$

So, $\mu = \dfrac{1}{1 + \dfrac{PROP}{\sum T_i}} = \dfrac{1}{1 + \dfrac{a}{N}}$

Note that as *N* increases there is less time wasted in propagating the token and utilization increases. As *a* increases the utilization decreases.

II. <3 points>

For token ring release after reception (RAR): $\mu = \dfrac{1}{1+a}$

For Ethernet (CSMA/CD): $\mu = \dfrac{1}{1+5a}$

For the same “a”, we see the $u_{\text{RAT}} > u_{\text{RAR}} > u_{\text{CSMA/CD}}$

When ‘a’ increases the utilization for all schemes drops, while at small ‘*a*’ they exhibit somewhat similar utilization.

III. <5 points> For CSMA/CD the number of probable collisions increases with the increase in *N* (the number of attached stations).
For token ring (release after reception), the number of stations sending per token-propagation around the ring increases, but so does the wait time (since each station would wait for its own transmission to be received), so the net result doesn’t change the utilization.
For token ring (release after transmission) the number of stations sending per token-propagation around the ring increases, so the wasted (idle) time decreases, which increases the overall utilization. (this also shows in the utilization formula).

Q2. (12 points: 3 + 6 + 3) MAC layer utilization
For an Ethernet LAN (shared bus) the data rate was increased from 10Mbps to 100Mbps.

- ***I.*** How will the utilization (U) of this network change? [Calculate U for each case]
- ***II.*** Suggest two ways in which we can return the utilization to what it was before (By increasing or decreasing another parameter and by how much? Show your reasoning.)
- ***III.*** One person argued that increasing the number of stations attached to the LAN would reduce the idle time on the LAN and hence increase the utilization. Do you agree?

A2.
I. For Ethernet U=1/(1+5a), where a = Tprop/Ttrans.
But Ttrans=#bits/data rate. Since the data rate was increased, so Ttrans is decreased (by a factor of 10), and hence a is increased by a factor of 10, and subsequently U is decreased.

II. We can either increase the # of bits per frame by a factor of 10 (so that Ttrans would decrease by 10), or decrease the length of the Ethernet network/cable by a factor of 10 (so that Tprop would increase by 10), or a combination thereof (i.e., increase the # of bits by 5 and the length by 2, so on).

III. No. Increasing the number of stations would increase the probability of collisions and so will bring the utilization down.

Q3. (10 points: 6 + 4)

I. What is the mechanism used to increase the efficiency of CSMA/CD for switched Gigabit Ethernet? Describe its operation and how it increases efficiency.

II. Discuss whether CSMA/CD is still needed for switched Ethernet and why? If not, why is it still being used?

A3.

I. The mechanism used is 'frame bursting' which allows a number of frames to be sent back-to-back without collision detection. In traditional Ethernet the sending station needs to wait until the frame propagates to the end of the LAN and back in order to perform collision detection, which increases the idle time and reduces utilization. By allowing a station to transmit a burst of frames without having to wait increases efficiency by reducing the wasted (wait) time.

II. CSMA/CD was designed to resolve contention and collision in shared media when multiple stations attempt to access the channel almost simultaneously. For switched Ethernet, there is no contention for the channel (since it is a point-to-point technology, similar to routers, where packets are stored/buffered in the switch, and a switching table is used). Hence CSMA/CD technically is not needed. However, for backward compatibility and 'interoperability' reasons, and to integrate well with legacy (already-installed) Ethernet systems, the same format of packets and same CSMA/CD mechanisms are used. This was one of the reasons why switched Ethernet was so successful (whereas ATM, another fast LAN technology was not).

Q4. (6 points) Can CSMA/CD be used in wireless networking? Why?

A4. In the wireless medium there is no direct correlation between the reception at the sender and the reception at the receiver which makes the collision detection by the sender inadequate to detect the collision at the receiver. So the answer is no, it cannot be used… more details are given in the lecture slides.

Q5. (6 points) What is the "hidden terminal" problem? and how does 802.11 address such a problem?

A5. In a chain topology of A-B-C-D, where A can hear only B and not C then when C is sending to B, A's carrier sense would not be enough to top collisions if it sends to B as well. (more details and scenarios/figures are given in the lecture slides)

To address the problem we use RTS/CTS exchange. RTS reserves the medium in the sender's neighborhood, and CTS reserves the medium in the receiver's neighborhood. This way, collisions can be avoided to a large extent, addressing the hidden terminal problem. (more details and given in the lecture slides/notes).

Q6. (8 points) "802.11 is fair always since it randomizes the timers before a station attempts to access the medium, so that the previous winners do not dominate the channel in the future", discuss this statement, whether it is true or not and when, illustrating your point with ample examples.

A6. In general, randomization of the timer takes care of situations where a successful node has several subsequent frames to send. However, the timers do not take care of situations where a node needs to defer due to ongoing communications in its neighborhood. An example would be a chain topology with A-B-C-D-E. If a transmission starts between A and B, so C has to 'defer'. Similarly, if a transmission starts between D and E then C has to 'defer'. In a scenario where A and B start, and before they end their transmission D and E start, and before they end their transmission A and B start again… so on, so forth, then C would starve (not have the chance to send any packets), which would results in '0' throughput for C leading to severe unfairness.

Different wording:
On one hand, the randomized timer setting does prevent previous winners (those that were able to capture the channel and send their packets in the previous contention cycle) from dominating the channel. However, this does not necessarily provide fairness in all scenarios. One example of severe unfairness can occur in a scenario with 5 nodes in a chain topology (say A-B-C-D-E) where A's neighborhood contains B, B's neighborhood contains A and C, C's neighborhood contains B and D, D's neighborhood contains C and E, and E's neighborhood contains D. In case B sends to A, then C needs to back off (i.e., enter 'defer' mode/state) unless the transmission is over. At the same time, in case D sends to E, then C needs to defer as well. A synchronized action in which B sends to A and before the transmission ends D sends to E, so on and so forth. This causes C to be denied access to the channel repeatedly, and unfairly.

Q7. (8 points: 4 x 2)

- ***I.*** What is 'rate adaptation' in wireless MAC protocols? Why is it used, and how does it help the performance of wireless networks?
- ***II.*** Discuss possible effects on upper layer protocols (such as TCP and real time applications)?

A7.

I. Rate adaptation in wireless MAC (e.g., 802.11) changes the encoding scheme to reduce the sending data rate when the signal to noise ratio SNR (or the signal power) decreases to a level at which BER is high (~10% or more). This could occur due to increased interference from the nearby environment or due to the increased distance from the base station (or the other node involved in the communication). By decreasing the data rate, the SNR is increased and the BER is decreased to acceptable levels. This effectively increases the usability (in terms of distance and operational range) of the wireless network.

II. Decreasing the data rate (sometimes by orders of magnitude; e.g., from 100Mbps to 1Mbps) would increase the delay dramatically for the wireless links. TCP relies on RTT estimates to detect congestion and set its retransmission timers. Having huge fluctuations in the RTT estimates (due to the rate adaptation of wireless links) can degrade the performance of TCP.

Q8. (15 points) Mention five advantages to using CDMA (or DSSS).

A8.

1- Resilience to fading and interference since the signal is spread in the frequency domain over a large spectrum (interference usually affects part of the spectrum and hence only a part of the spread signal)
2- Security: only the intended recipient (that has the right seed for the pseudo noise (PN) sequence) is able to decipher/decode the message, the other recipients will not be able to receive it and it would look like noise to them
3- No frequency planning since the users use the same frequency
4- Soft handoff can simplify the process of hand off during mobility
5- The number of users supported is no longer limited to the available spectrum, since the separation between users occurs in the code (not the frequency or time) domain. Typically a CDMA system can support 3-5 times the users in a TDMA system. The max capacity is actually limited by the noise floor (since to a user all the signals not destined to it would result in 'noise').

Q9. (24 points: 3 + 9 + 6 + 6) TCP over ATM:

I. What is the main observation (or problem) you can identify regarding transfer of TCP segments and IP datagrams over ATM networks?
II. Suggest and compare three different schemes that attempt to alleviate the above problem. Mention the advantages and disadvantages of each.
III. What are the features in ATM AAL5 that allow efficient transfer of TCP over ATM? Mention 2 main features and explain how they can be used to increase the efficiency (define what you mean by efficiency).
IV. What would happen if those two features were not supported?

A9. I. -observation: when ATM cell is dropped, all other ATM cells that belong to the same IP datagram are useless

II. -solution: develop discard strategy to minimize transmission of useless cells
(1) Partial Packet Discard (PPD):
-when a cell is dropped at a switch, all cells belonging to the same datagram are dropped
-switch identifies the end of IP datagram using type-bit in ATM header in AAL 5
-on average: ½ datagram worth of ATM cells are transmitted uselessly

(2) Early Packet Discard (EPD):
-when buffer exceeds a threshold, drop complete IP datragrams
-problem of fairness: the shorter the datagram, the higher the probability of drop

(3) add fairness using fair buffer allocation (FBA):

- when EPD is invoked drop from connections using more than their fair share
-the number of VC connections is V
-if N is the current occupancy, then the fair share is N/V
-the weight w(i)=N(i)/[N/V], where N(i) is occupancy of connection I
-policy to drop: if (N>R) and w(i)>z then drop, where R is the congestion threshold and z~1

III. The features: 1. The 'type bit' is set to '1' in the very last cell that belongs to a datagram. This helps delimit the datagrams at the ATM layer.
2. The virtual channels (VCs) are not multiplexed, so that every VC at the ATM layer corresponds to a flow at the IP layer. This ensures that consecutive ATM cells belong to the same IP flow.
Both of these features are necessary to allow the packet discard mechanisms (discussed in Q8) to work effectively.

IV. if those features are not supported, then ATM will not be able to identify the beginning and end of each datagram and the packet discard mechanisms would not work effectively. Hence, the efficiency of the ATM cloud would decrease.

Q10. (16 points: 4 + 6 + 6)

I. What are the main problems that Mobile IP is attempting to solve and how does it solve it?
II. What are the problems with Mobile IP?
III. Suggest two different ways to overcome the shortcomings of Mobile IP by discussing other mobility management solutions

A10.
I. Mobile IP strives to provide (1) "transparency" of connection while providing the ability to move for mobile nodes. This means that the nodes communicating with a mobile host do not know its current location, nonetheless they are able to establish continuous communication with it.
(2) maintaining established connections (such as TCP) during mobility.

Mobile IP provides this transparency by having the correspondent node (CN) send packets to the home network, then the home agent (HA) would tunnel these packets to the mobile node. The home agent knows the (new) address of the mobile node through registration process that occurs through the foreign agent (FA) (in the foreign network).

II. The main drawbacks of basic 'Mobile IP' are: (. 3 drawbacks are sufficient)

1. Triangle routing 2. Delay (or packet loss) during handoff 3. Communication overhead during handoff 4. Not fit for micro mobility

III. Two suggestion to deal with the triangular routing problem are:

- route optimization: the mobile node sends its new address to the correspondent host so that further packets are sent directly to the mobile node without going to the home agent. [or the home agent sends this new address to the correspondent host]
- multicast-based mobility: have the mobile node use a multicast address (which is independent of location), and whenever it moves to a new network it joins that multicast group through the new network and leaves it through the old network. The correspondent node always sends to the group.

Extra:
The advantage of both the above proposals is that they avoid triangular routing. In the multicast-based mobility approach, there is no need for home agent or foreign agent, and the handoff is theoretically simpler and faster than other approaches.

Disadvantages:
Route optimization: may incur handoff problems as does Mobile IP. Needs the correspondent node to change address of the mobile host on the fly for the same connection (like TCP connections).

Multicast-based mobility: needs to assign a multicast address for each mobile node. Connection needs to be established with a multicast address, which may cause a problem for come applications (like TCP).

Q11. (10 points: 3 + 3 + 4)

I. What is the spectral congestion problem is mobile networks?

II. How did cellular networks overcome such a problem?

III. What are the main problems created by the solution in *II.* above? [mention two main problems at least]

A11. (more details are given in the lecture)

I. Spectral congestion: Limited spectrum to provide communications, which limits the number of channels and subsequently the number of users. The contention of users for the limited channels/spectrum causes the spectral congestion

II. Instead of using a single powerful tower to cover a large area with a limited number of channels, smaller towers are used with limited geographic coverage. This allows for re-use of the spectrum in disjoint geographic areas, and hence the support of more number of users over a larger geographic extent.

III. The two main problems include: 1- handoff: since the extent of every cell is smaller than before, then the number of transitions between them through a trip (e.g., using a vehicle) will increase, which would require frequency handoff between the cells, and 2- frequency planning to combat interference (both adjacent channel and co-channel intereferences)… more details given in the lecture.

Q12. (10 points: extra) In a wireless ad hoc network, where all the devices are mobile and all hosts are acting as relays (or routers), do you think the unicast routing approach in the Internet is suitable? Why? If not, suggest and discuss an approach that may work in a better way. [Assume that the ad hoc network is always connected and that disconnections are rare].

A12.
In a mobile environment the topology is highly dynamics and the paths may be constantly changing. A proactive approach, like that used in the Internet, would calculate the routes using routing protocols in anticipation of packets. However, in ad hoc networks the calculated routes become invalid quickly, and so there is a lot of control wastage in calculating routes that may never be used. The proactive (or table driven approach) is hence inadequate for ad hoc networks.

A reactive approach may perform better, where the protocol only discovers and calculates routers for those flows that are active. When a flow is requested, a route discovery phase is invoked to discovery valid routes for that flow, which may incur a startup delay, but generally responds much better to network failures and dynamics than a proactive approach.

Q13. (10 points: extra) In delay tolerant networks, DTNs, the network has mobile devices (as in ad hoc networks) but they get frequent disconnections where the network is partitioned, and there may be times where there are no complete routes from any point to any other point. Suggest another paradigm of routing that would be suitable for such an environment. Discuss briefly how to enhance the efficiency of your mechanism.

A13.
In DTNs, a path may not be valid or complete at any one point in time, but paths may be created over time. This brings up the issue of routing in time and space domains. One way to achieve such a goal is store information in intermediate nodes, and after moving and encountering other nodes, the message can be then forwarded to other encountered nodes.

One simple way to route in such an environment is to ‘flood’ or use epidemic routing where a message is given to every encountered node. This maximizes the probability of message delivery and minimizes the delivery delay, but adds significant overhead to the network.
On the other hand, having random delivery alleviates the problem of overhead but at the expense of increased delay and reduced delivery rate.

A method to guide messages to the destination, perhaps based on location information, or previous encounter patterns may result in reduction in overhead while attempting to maintain the success/delivery ratio and delay.