Leeson Chen
Homework 3 for CNT 4007C Computer Networks Fundamentals
November 13, Fall 2019

On the Data link layer, MAC protocols and Wireless Mobile Networking Due Date: Nov 14th, 2019 on Canvas *before* midday 11:59 AM. Please also provide an exact hard copy (in class or during the TA office hours).

Q1. Error detection and correction:
I.Describe how parity bit techniques can *detect and correct* errors in data-link frames? What are its limitations?
A single parity bit will determine if the number of 1s in data is even or odd. and thus can detect if a single bit has been corrupted. Similarly, two dimensional parity does this for all rows and columns of a data matrix. However, if two bits have been flipped it is harder to detect, since there will be an even or odd number as it is supposed to be. An incorrect parity will cause the data to be detected as corrupt, and resent therefore correcting the error.

II. Provide an example where the error cannot be detected or corrected. Explain why.
As staed above, flipping two bits results in the even number still being even, and a parity bit will not detect that.

III. Provide an example where *multiple* errors can be detected and corrected by parity techniques. Explain the limit of your example.
Multiple errors could be detected by a two dimensional parity matrix, such that even if a singular row or column goes undetected, the perpendicular rows or columns will have an odd number of errors and thus detect the corrupted bit. However, this is limited to specific corruptions, and only for data formatted in a 2d matrix.

Q2. Provide analogy between the operation of CRC and CDMA in terms of how they function at the sender and receiver.
CRC stands for cyclic redundancy check, and is a type of error detection used in Ethernet frames, wifi, and ATMs. The function is sending D data bits and R CRC bits so that dividing D by G results in R after performing it 2^r times. An analogy would be as if you mailed a friend some seeds, and they had to plant and water them, and determine what they grew into. If they grew into the correct plant they'd know the seeds were what you meant to send, and if they were a different plant or didn't grow they would be wrong. CDMA is a wireless multiple access protocol for code division, where soft handoff is enabled by not changing the channel or frequency of cells. This is akin to texting your friend, but then he moves from his phone to his laptop and then uses facebook messenger to talk to you, then moves to his work computer and talks to you through email. You're only changing the base station, but still getting the same information.

Q3. Someone suggested to get rid of the MAC address. 'You only need IP addresses' to route packets! Is this a true statement? Discuss the main implications and consequences of this suggestion on LANs, switching operations (e.g., in data centers), and mobility.

Getting rid of MAC addresses would not help routing packets, as MAC addresses and IP addresses have very different characteristics. For example, MAC addresses are fixed to a device, whereas IPs can change. MAC addresses are an Ethernet standard, while IP addresses are an Internet standard. IP addresses are for the most part known around the world, while MAC addresses aren't. A MAC address provides more distinction between devices than IP, which may overlap between devices (for example, a virtual machine). MAC addresses are more helpful for the physical layer of hardware, and within a LAN. If MAC addresses were to be removed and replaced with IP only, then within a LAN you would need to assign IPs while avoiding collisions with existing IP addresses, and keep track of users that move (which is difficult when IP routing tables are based on location).

Q4. Main communication paradigm over shared media:

I. Describe how communication on a shared medium (e.g., Ethernet bus/LAN, or token ring) occurs on the wire at the MAC (data-link) layer for unicast, broadcast and multicast.

When broadcasting, the sender essentially ignores MAC addresses and just floods the LAN regardless of destination. This is used for the self learning tables discussed later. For multi and unicast, the specific MAC addresses of the destination devices is used.

II. Describe the vulnerability of such method with potential attacks.

One problem with the self learning and forwarding tables is that it requires the destination to confirm it is the correct receiver. A malicious device could pretend that it is the intended recipient, thus registering itself in the table as the destination. Then all future data intended for the recipient would instead be sent to the bad actor.

Q5. Utilization of MAC protocols

I. Derive the utilization formula for:

(a) Ethernet,

Utilization = 1 / (1+5a) where a = Tprop / Ttrans —> average wasted time = 5xTprop

(b) Token ring (release after reception), and

Utilization = Tuseful / Ttotal = Tuseful / (Tuseful + Twasted) = 1 / (1+a) where a = Tprop / Ttrans

(c) FDDI [which uses Token ring (release after transmission)].

Utilization = 1 / (1 + (a/N)) where a = Tprop / Ttrans, and N = # of stations (better)

II. Compare them.

Ethernet has the lowest utilization for the same value of a. RAT FDDI is better than RAR token ring. Performs better for a lower number of stations than RAR.

III. Comment on what happens when the number of stations on the LAN is increased in each case.
For token ring, if the number of stations N increases, the overall utilization increases because less time is spent token passing and waiting. Release after transmission (RAT) is even better than release after reception because within the same token rotation time more stations can send.

Q6. MAC 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]
The utilization decreases because while

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.)
Increasing Tprop proportional to Ttrans will result in the ratio a remaining the same, and thus the utilization will return to what it was before.

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?
This statement is true for specific scenarios, such as token ring release after reception, where more stations reduces rotation time and thus increases utilization. However, it is not true for all scenarios, because more stations on an Ethernet bus would result in more collisions, thus reducing utilization.

Q7. Discuss whether CSMA/CD is still needed for switched Ethernet and why? If not, why is it still being used?
Ethernet alone is unreliable and connectionless. It has no handshaking, ACKs, or NACKs built in for the sending and receiving NICs to communicate. However, Ethernet integrated with the CSMA/CD algorithm allows for many features such as collision detection and exponential backoff. The algorithm is highly efficient, with better performance than ALOHA while being cheap and decentralized. Switch Ethernet uses frame bursting, which increases utilization while still being CSMA/CD compatible. While CSMA/CD may not be completely necessary for switched Ethernet to function, the benefits of keeping it integrated outweight the benefits of removing it.

Q8. Why do we need mapping between IP address and MAC address? What table provides such mapping?
The fastest way to determine an interface's MAC address when the IP addres is known is by creating a map. These maps are provided by ARP tables, where ARP stands for address resolution protocol. Every IP node on a LAN will contain a copy of the table. A single table entry consists of the LAN node's IP address, MAC address, and TTL (time to live before updating).

Q9. Switched ethernet: Which data structure does an ethernet switch use to avoid the inefficiencies of a hub? How does it construct such data structure?
In order to avoid a hub's inefficiencies, an ethernet switch constructs a map of every MAC address to its corresponding interface. This map can be thought of as a hash table, in terms of data structures. To create the map, a switch uses self-learning and forwarding, where if it doesn't know a mapping it floods to find the destination, then updates the map. Sometimes map data is out of date or incorrect (assigned a time-to-live, or TTL), in which case the switch will flood again to get new mappings.

Q10. Data center networks:
I. What are the two main functions of a data center network?
Data centers are buildings full of thousands of server hosts, used by large tech corporations such as Google, Amazon, and Youtube. The problem is that all these servers are running discrete processes and contain different data, but requests from all around the Internet are requesting service—whether for computing power, lookup, indexing, viewing content, or other various functions that these centers provide to clients. On top of serving thousands to millions of clients, the data center itself must cope with managing and balancing load, avoiding processing, networking, and data bottlenecks.

II. What are the two main functions of the data center load balancer (sometimes referred to as a 'layer 4' switch)?
Above every rack of servers is a top-of-rack switch, above which are several more layers of switches connecting those switches, and culminating in a load balancer / layer 4 switch. The load balancer's functions are to detect which servers are doing heavy vs light computational work, and to balance incoming client requests from the Internet within the center most efficiently. The client would then receive the result, without having known the internal details of the data center.

Q11. Provide four potential example use cases of infrastructureless multi-hop networks (such as adhoc networks and peer-to-peer wireless networks).
A wireless ad hoc network requires no infrastructure except for the hosts participating. This would be useful in areas where infrastructure does not exist yet, cannot exist, or has been taken down or damaged. Some examples would be the communication within parties of protestors, search parties in the wilderness, military operations in foreign territory, or between ships on the open ocean.

Q12. What are the main differences between wired and wireless networks (mention 3 main reasons) and further elaborate on one of those reasons in three sub-points.
A wireless network offers the user mobility: that is, the user can roam around in the vicinity of a wireless access point with their phone or laptop. The base station is often a cell tower or 802.11 AP (wifi router). The wireless end hosts are often restricted by battery power, meaning that Internet processes cannot consume high amounts of endless energy. And wireless links are more prone to error and weakness, called bit

error rate (BER). Signal strength can decrease as a function of distance from base station, but also comes from interference with other wireless devices, and multipath propagation.

Q13. Wireless rate adaptation:

I. What is rate adaptation in 802.11, and why is it used?

In 802.11, a base station and its connected mobile can work together to perform rate adaptation, where the transmission rate is dynamically changed when the mobile moves. This varies the SNR (sound-noise ratio): as SNR decreases, the BER (bit error rate) increases if the node moves further from the base station. In short, when a signal is strong a high data rate is used; when the signal is weak a lower data rate is used. Rate adaptation is used to optimize different data rates for their corresponding signal strength, to get the most data rate out of every signal but not overwhelm or congest the network and mobile.

II. Discuss in detail its effect on TCP operation and dynamics.

Because two main features of TCP are congestion control / avoidance and ensuring packets are sent in order, rate adaptation helps these features exist in wireless connections. By dynamically changing the rates, a weak signal will not be overloaded with a high number of packets (which would cause congestion). Similarly, a weak signal receiving more data than it could handle would start to drop packets, so rate adaptation avoids this as well.

Q14. Hidden terminal:

I. What is the hidden terminal problem?

The hidden terminal problem is an issue in wireless systems such as 802.11, where a node A can communicate with another wireless access point, node, or terminal B, but not a third node C which B is connected to.

II. Why does it not exist in wired networks (like Ethernet)?

The underlying cause of the hidden terminal problem is signal attenuation, caused by interference of signals within the same frequency drowning each other out. This is directly due to the mechanisms of wireless networks, whereas a wired network such as Ethernet would not use wireless signals, and thus does not suffer from the hidden terminal problem.

Q15. What type of interference is directly reduced by frequency planning in cellular networks? and how is it reduced?

Cellular networks without frequency planning are prone to the problems of frequency limitation, spectral congestion, and user capacity. This comes from many cells within the same area using the same frequencies, which will become overloaded. To solve this, frequency planning maximizes distance between like-frequency cells using a tesselation pattern.

Q16. What are the dimensions in which we can separate the signal for the purpose of multiplexing? Mention five, not including hybrid!
Multiple access can be achieved by splitting the signal through its frequency (FDMA), time (TDMA), space (SDMA), or code (SSMA)—which itself can be subdivided into frequency hopping (FHMA) and code division / direct sequencing (CDMA / DSMA). The hybridization of these five techniques allows for even more possibilities.

Q17. What are the advantages of using spread spectrum (e.g., CDMA) over other multiple access techniques we have studied? Mention five advantages, and for each explain how spread spectrum achieves its advantage.
The underlying mechanism of CDMA spread spectrum is that every user's signal is encoded with a pseudonoise (PN) sequence orthogonal to the data. This brings the following advantages:

- Security: the signal can only be decoded if the other side also has the PN sequence. Any other decoding results in random noise, thus the transmission is securely encrypted.
- Coexisting: multiple users can transmit at the same time, without interfering with one another's signals, provided the codes themselves are orthogonal. The number of users is not limited by the spectrum, but the noise floor, which allows for approximately 15x more users than FDMA, and 5x more than TDMA.
- Fading effects are reduced: signal power is spread over a wide frequency spectrum
- No frequency planning: the PN encryption allows all users to share frequencies
- Soft handoff: since all users share the same frequency, just the base station changes for handoff

Q18. Someone suggested that 'CDMA systems are limitless in terms of number of simultaneous users supported', argue (with details) for or against this statement.
This is not a true statement, because while a theoretically infinite number of users can continue joining the spectrum, each one will require their own PN sequence, and as the number of PNs goes up the noise for each signal when decrypted also goes up. Effectively, the limit of simultaneous users is reached when there is too much noise for any individual user to hear past the noise of other signals. This effect is called the noise floor.