Lecture 2
MAC II: Collision Avoidance and Controlled Random Access

Reading:

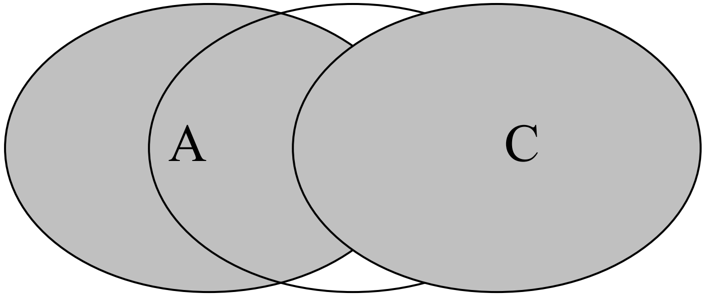
- V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A Media Access Protocol for Wireless LAN's," *Proc. SIGCOMM '94*, September 1994, pp. 212-225.
- D. Goodman, R. Valenzuela, K. Gayliard, and B. Ramamurthi, "Packet Reservation Multiple Access for Local Wireless Communications," *IEEE Transactions on Communications*, Vol. 37, No. 8, August 1989, pp. 885-890.
- J.-C. Chen, K. Sivalingam, P. Agrawal, and S. Kishore, "A Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption," *Proceedings of IEEE INFOCOM '98*, April 1998.

Location-Dependent Carrier Sensing

- CS depends on location of the node
  - Transmitter performs CS
  - Know the state of the channel at transmitter
  - Cannot determine the state of the channel at the receiver
- Causes three problems
  - Hidden Nodes
  - Exposed Nodes
  - Capture

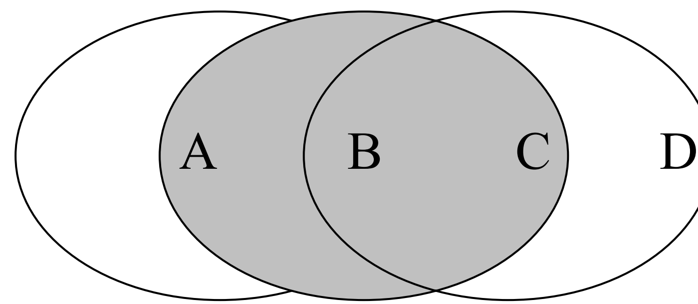
Hidden Terminal Problem

- A node that is within range of the receiver but not within range of the transmitter
- A hidden node will sense an idle channel and transmit data
  - Suppose node A is transmitting to node B
  - Node C will sense an idle channel
    - A is outside the range of node C
    - C will transmit data to node B
  - Causes collision
- Problem: collisions must be detected at the transmitter but actually occur at the receiver
- Nodes outside the range of the transmitter "hidden" to the transmitter → CSMA not effective



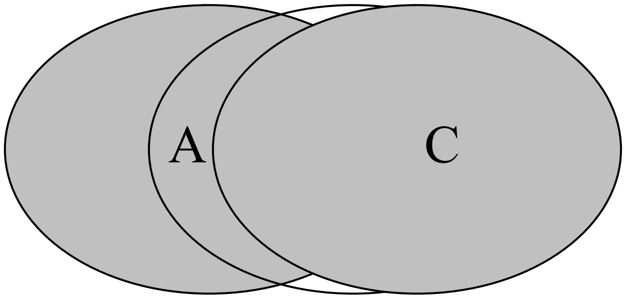
Exposed Nodes

- A node that is within range of the sender but out of range of the destination
- An exposed node will sense a busy channel and not transmit
  - Suppose node B is transmitting to node A
  - Node C will sense a busy channel
  - However, node C could transmit to node D
    - D is outside the range of node B
    - No collision
  - Node C (exposed node) backs-off
- Causes underutilization of the channel



Capture Effect

- When receiver can cleanly receive a transmission from one of two simultaneous transmissions
  - Collisions may not cause both packets to be "lost"
  - Strongest user can successfully capture receiver
  - Packet from strongest user may survive collisions
- Often closest user captures receiver due to small propagation path loss



Node C captures the channel from Node A

Capture Effect (cont.)

- Minimum power ratio of an arriving packet relative to other colliding packets such that arriving packet received = capture ratio
- Capture ratio depends on receiver and mod. method used
  - Minimum power ratio which one arriving packet must have relative to the other colliding packets in order that it can be received successfully

Capture Effect (cont.)

- If capture is included in throughput analysis, MAC protocols can achieve higher throughput
  - Throughput analysis typically assumes no capture and is a lower bound on throughput in practice
  - Upper bound on throughput can approach one if some station always captures the colliding packets
  - True throughput lies somewhere between two bounds but closer to lower bound
- Capture results in unfair sharing of the bandwidth
  - Near-far effect: Nodes close to the receiver can easily capture the receiver and prevent nodes further away from communicating with the receiver
  - Wireless MACs aim to ensure fairness in the presence of capture

CSMA/CA: CSMA with Collision Avoidance (CA)

- Collision avoidance implemented using ACKs, backoffs, packet reservation
- CA implemented using
  - Out-of-band signaling
    - Receiving nodes transmit a signal to let nodes in their range know that the channel is busy
    - Eliminates hidden nodes
    - Increases the number of exposed nodes
    - BTMA, DSMA, etc.
  - Control handshaking
    - Each stage of the handshake lets nodes know if they are in range of the transmitter, receiver, or both
    - Tradeoff in overhead for handshaking and the number of hidden nodes eliminated
    - MACA, MACAW, etc.

CSMA with Busy-Tone Signaling (BTMA)

- CSMA with Busy-Tone Signaling (BTMA)
  - Whenever node receiving message, can send a signal on the "busy-tone" channel to let all nodes in communication range know that the medium is busy
  - In centralized network, BS sends busy-tone signal
  - In decentralized network, all nodes must transmit busy-tone signals when they detect busy channel

Data Sense Multiple Access (DSMA)

- Forward control channel used to inform nodes if reverse channel is idle or busy
- Nodes detect a busy-idle msg on forward control channel
- If busy-idle message indicates no users transmitting on reverse channel, a user can send its packet
- Does not eliminate collisions
  - Multiple nodes might note free channel and send data simultaneously → collision
  - Retransmission strategy as in CSMA
  - Similar to BTMA but busy-tone information piggy-backed on control information being sent from BS to mobiles

MACA

- Problem with CS is that transmitter cannot determine state of carrier at receiver
  - Hidden terminal problem
  - Exposed terminal problem
- MACA implements hand-shaking collision avoidance
  - Precede data transmission with request-to-send (RTS) packet
    - RTS contains length of expected data transmission (all phases)
    - All nodes in vicinity of Tx node enter backoff for duration of message delivery
  - If Rx node successfully receives RTS, replies with a clear-to-send (CTS) packet
    - CTS packet contains length of expected data transmission
    - All nodes in vicinity of Rx node enter backoff for duration of message delivery
  - Upon receipt of CTS, Tx node sends data

MACA (cont.)

- MACA successfully alleviates the hidden terminal and exposed terminal problems
  - All nodes that could corrupt transmission (any nodes hearing CTS) know to not transmit – virtual carrier sense
  - All nodes that would not corrupt transmission (any nodes only hearing RTS and not CTS) know can transmit without corrupting communication between Rx and Tx
- MACA assumes transport layer reliability (e.g., TCP)

MACA (cont.)

- Backoff
  - If a node does not receive a CTS in response to an RTS, it will eventually assume a collision occurred
  - Other nodes in vicinity also trying to use the channel
  - Node backs off before retransmitting RTS to reduce probability of another collision
    - Node chooses a random number x between 1 and BO
    - Node must wait x time slots before retransmitting
  - MACA uses binary exponential backoff (BEB)
    - If an RTS is unsuccessful (i.e., it does not elicit a CTS), the BO is doubled (up to a maximum value)
    - Otherwise, wait time set to minimum value

MACA (cont.)

- Collisions usually occur in RTS/CTS packets
  - If RTS/CTS packets successful, most nodes know that channel is busy and will defer their own transmissions
    - Some nodes may not receiver RTS or CTS due to other transmissions in their vicinity
  - Data has a high probability of no collision
- RTS/CTS packets short compared to data packets
  - Packets destined to collide do not use much of the channel resources
  - Packets that do require large amounts of channel resources have a high probability of being successfully transmitted
  - Good policy!

MACAW

- MACAW adds functionality to ensure reliable delivery of data at the link layer
  - Suppose node A sending data to node B
  - Node A sends RTS
  - Node B sends CTS
  - Node A sends Data
  - If data received successfully, node B sends ACK

MACAW (cont.)

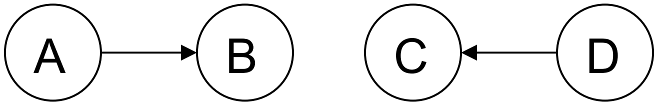
- If node A does not receive ACK after a certain time, node A retransmits RTS
- If node B had received the data
  - ACK message was lost, not data message
  - Node B sends ACK in response to RTS
- If node B did not receive the data
  - Node B sends CTS in response to RTS
  - Data is retransmitted
- Node A’s backoff counter increases if initial data transfer was unsuccessful to minimize future collisions

DS Messages in MACAW

- Nodes in range of transmitter (A) but not receiver (B) (e.g., exposed nodes) may not hear CTS and thus not know if RTS-CTS exchange successful
- If these nodes backoff until end of expected data transfer, might be a waste if CTS not received by A
- Nodes can employ CS when expect node A to be tx data
  - If channel sensed idle, nodes can assume RTS-CTS unsuccessful and node A is currently backed-off
  - If channel busy, nodes must defer their own transmission
- Instead of CS, MACAW uses Data Send (DS) msgs
  - Transmitter (A) sends DS message after receiving CTS to inform exposed nodes that A will now send data packet
  - Nodes receiving DS must defer their tx for the allotted time
- DS messages ensure fairness

RRTS Messages in MACAW

- RRTS messages
  - Possible to "starve" nodes from channel access
  - A transmitting to B, D transmitting to C
  - If A is sending data when C receives D's RTS, C cannot send a CTS due to required deferral
  - D will backoff
  - Since data messages are long, D may backoff to maximum backoff value
  - A will continuously be able to transmit successful RTS-CTS messages and capture the channel



RRTS Messages (cont.)

- D must be lucky enough that the random time chosen to transmit an RTS falls during the time between an ACK and RTS from the A-B stream
- Use request-RTS (RRTS) messages
  - When A-B data transfer complete, C sends an RRTS packet to D
  - When D receives RRTS packet, D sends RTS
  - D-C stream can now commence
  - Ensures fairness

Backoff in MACAW

- BEB can result in one stream starving another
  - With every unsuccessful RTS, node will BO
  - Node with successful RTS will have BO at minimum value
  - Node with unsuccessful RTS will have BO at maximum value
  - If no max BO, one node would eventually have infinite BO
    - Other node would permanently capture channel
    - High throughput but not fair allocation
- All pkt headers modified to include node's BO value
  - Nodes receiving packet copy this BO value
  - All nodes will have equal BO values
- BO algorithm modified to mult inc, linear dec
  - Upon collision, BO increased to min(1.5*BO, BO_{max})
  - Upon success, BO decreased to max(BO-1, BO_{min})

Distributed Foundation Wireless MAC (DFWMAC)

- 4-way handshake: RTS-CTS-DATA-ACK
- Node with data sets a timer for a random time
- Timer decremented while channel is idle
- When timer expires, node sends an RTS
- Receiving node sends CTS if RTS received successfully
- Sender sends data to receiver
- Receiving node responds with an ACK
- If ACK not received, packet assumed lost and sender retransmits the packet
- If the RTS fails, node backs-off using a BEB policy
- Preference given to sending ACK over RTS
  - Node must wait at least DIFS interval before attempting RTS
  - Node must wait at least SIFS (< DIFS) interval before attempting ACK

DFWMAC (cont.)

- DFWMAC uses CS before handshaking
- Virtual CS in addition to physical CS
  - Each packet has time fields that determine expected length of entire data exchange sequence
  - RTS time field includes time for CTS, Data, and ACK packets
  - CTS time field includes time for Data and ACK packets
  - Data time field includes time for ACK
    - ACK time field set to 0 to signify end of the message exchange
  - Nodes receiving any of these packets set their Network Allocation Vector (NAV) field
    - Nodes are not allowed to access the channel for NAV amount of time
    - All nodes wait until their NAV timer expires before they begin sensing the channel
- DFWMAC used as the IEEE 802.11 MAC protocol

Controlled Random Access MAC

- RA methods can have low throughput due to collisions
- When tx begun, full time and frequency resources of channel are being used even though sender cannot be certain transmitted packet will not encounter collision
  - Packet destined to collide → channel resources wasted during transmission of these packets
  - Requires more control to better utilize resources

Reservation-ALOHA

- Reservation-ALOHA (R-ALOHA)
  - Use ALOHA contention to reserve slots, TDMA for data tx
  - Time axis divided into frames, frames divided in slots
  - Some slots subdivided into reservation subslots
  - In **unreserved** mode, each user with data sends a reservation request in one of reservation subslots
  - Intended recipient sends ACK and slot assignment if reservation request does not suffer collision
  - In **reserved** mode, one slot assigned to hold reserv. subslots, all other slots pre-reserved data tx slots

Reservation-ALOHA (cont.)

- Node that has been granted slot sends its data in reserved slot
- Reservation exchanges heard by all nodes
  - Nodes know which slots are reserved and not to send data during these slots
- Flexible-TDMA
  - Contention-based reservation exchanges confined to short reservation subslots
  - Message slots shared among nodes with data to send in noninterfering manner
  - Design tradeoff of how many reservation slots and how many data slots
  - Control distributed among all nodes in the network → each receiving node grants reservations when tx node requests a slot

Polling Techniques

- Controller polls each node to see if has data to send
  - If node has data to send, it sends data to controller after being polled
  - If not, node sends negative reply (or nothing) and controller polls the next node
- Requires centralized control
- High overhead if nodes do not have data to send
- Requires controller to know about all nodes in the network

MAC for Different Traffic Types

- Bursty, short messages
  - Contention protocols
- Bursty, long messages, large number of users
  - Reservation protocols
- Bursty, long messages, small number of users
  - Reservation protocols with fixed TDMA reservation channel
- Stream or deterministic (voice)
  - Fixed access

Integration of Voice and Data

- Objectives of 3G systems: use single wireless system for all types of communication (voice, data, image/video)
  - MAC that is efficient for voice not efficient for data
- Contention-based protocols (e.g., ALOHA, CSMA)
  - Best for data communication
  - Best for large number of users with bursty traffic
  - No centralized control required
  - Can support variable number of users
  - Become inefficient when traffic load is heavy
  - Unpredictable and possibly excessive time delays not suited for real-time applications (e.g., voice)
- Fixed access protocols (e.g., FDMA, TDMA, CDMA)
  - Ensure fixed time delays needed for real-time applications
  - Lack flexibility for efficient integration of multiple services

Integration of Voice and Data (cont.)

- Voice application requirements
  - Packets can tolerate errors and packet loss (< 1-2%) without degrading service
  - Real-time constraints
  - Fixed assignment (FA) protocols most appropriate
- Data application requirements
  - Packets can tolerate delays
  - Cannot tolerate errors or lost packets
  - Random access (RA) protocols most appropriate
- Can separate frequencies into 2 bands
  - One band for isochronous packets (packets with real-time constraints)
  - Other band for asynchronous packets

Integration of Voice and Data (cont.)

- Obtain more efficient system by taking advantage of voice activity factor to send data during silent periods (~60% of typical voice conversations)
  - Data users can transmit during free TDMA time slots or free FDMA frequency channels as they become available
  - In CDPD, data users simply transmit in frequency bands not currently being used by mobile telephone users
    - Eventually will take advantage of "talk spurts" to increase data communication capabilities
  - In CDMA, each user has access to all BW-time resources
    - Resource to be managed is signal power
    - Integration of data communications with voice communications is straightforward: each user (either voice or data) uses a unique code to access the channel

Integration of Voice and Data using TDMA

- Movable boundary TDMA with silence detection
  - Frame divided into two regions with movable boundary between them
  - First region used for both voice and data traffic, but voice traffic has priority
  - If empty slots in first region (not enough voice traffic to fill first region), remaining slots used for data traffic
  - Second region reserved for data traffic
  - Boundary between voice and data regions moves depending on number of active voice packets in each frame
  - Max number of voice packets per frame fixed to ensure some minimum data traffic capacity yet ensure blocked voice packets kept to some minimum value

Packet Reservation Multiple Access (PRMA)

- MAC for variable mixture of voice and data packets
- For use in centralized networks
- Time segmented into frames
  - Frames contain fixed number of time slots
  - Frame rate = arrival rate of speech packets
- Information either **periodic** or **burst**
  - Periodic: speech, data if it is a long stream of information
  - Burst: short data packets
  - A bit in the header of the packet specifies what type of packet it is

PRMA (cont.)

- Nodes identify each slot as **reserved** or **available** based on feedback message received from BS
  - A reserved slot can be used only by the user terminal that reserved the slot
  - An available slot can be used by any terminal not holding a reservation that has information to transmit

PRMA (cont.)

- If a node has periodic information to send
  - Start transmitting in contention for the next available time slot (using S-ALOHA)
  - When the BS detects this first packet successfully, BS grants node a reservation for exclusive use of the same time slot in the next frame (via ACK msg)
    - Node now "owns" that time slot in all succeeding frames as long as it has an unbroken stream of packets to send
    - After the end of the information burst, node sends nothing in its reserved slot
    - BS transmits a NACK feedback message indicating that the slot is once again available
  - If first packet unsuccessful, node retransmits packet with probability q in subsequent unreserved slots

PRMA (cont.)

- To contend for unreserved slot, node must have permission
  - Determined by permission probability
  - Permissions of nodes independent
- Node attempts to transmit initial packet until BS ACKs successful reception of the initial packet or until packet is discarded by the node b/c it has taken too long to transmit the data
  - D_{max} = maximum holding time for a packet
  - Design parameter of the PRMA system
  - Must be determined by delay constraints on speech communication
  - If node drops first packet, continues to contend for a reservation to send subsequent packets

PRMA (cont.)

- Node drops additional packets as they exceed D_{max}
  - FIFO buffer ensures all packet losses occur at the beginnings of talkspurts
  - Less objectionable to listeners than random packet losses
- Data can tolerate large delay, so D_{max} can be set to ∞
- If a node has burst data to send
  - Node transmits data in an unreserved slot
  - If collision occurs, packet retransmitted with prob. r
- If $r < q$, system gives priority to periodic information

Power Consumption of MAC Protocols

- Radio operates in 3 modes: transmit, receive, standby
- Relative powers
  - $P_{TX} > P_{RX} >> P_{SB}$ for long-range communication
  - $P_{TX} \sim P_{RX} > P_{SB}$ for short-range, low power transceivers
  - Different MAC protocols will be "low-power" depending on model of transceiver power dissipation
  - Time delay and power dissipation switching between states
- To minimize energy dissipation
  - Place nodes in standby mode as much as possible
    - Nodes do not need to be on when not receiving data
    - Requires nodes to know when they must listen to the channel and when they can "sleep"
    - MAC protocols cannot use "promiscuous" mode to listen to other conversations
    - Node must know when other nodes have data to tx to it

Power Consumption (cont.)

- Tradeoff energy consumption and delay in receiving a message
- Approaches
  - Directory approach: BS broadcasts directory of packets waiting in its queue
    - Node receives directory and knows when to wake up and listen for data
  - Grouped-TDMA approach: nodes grouped and each group wakes up at given slot to determine if data needs to be received
  - Pseudo-random approach: nodes have unique pseudo-random sleep/wake cycles known to BS

Power Consumption (cont.)

- Collisions should be minimized
  - Retransmissions expend energy
  - Introduce delays
  - Reduce number of ACKs required
  - Use contention for reservations and contention-free for data transmission
- Allocate contiguous slots for transmission/reception
  - Avoids power/time in switching from Tx to Rx
- Have the BS buffer packets for a node and transmit all packets at once
  - Allows node to remain asleep for long time
  - Trade-off in delay to receive packets and BS buffer size

Power Consumption (cont.)

- Centralized scheduling is most energy-efficient
- Energy advantages depend on relative power in the transmit and receive mode