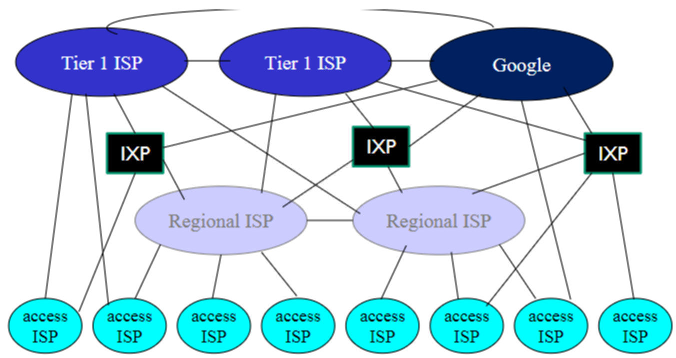
Lecture 2 Network architecture and work flow.

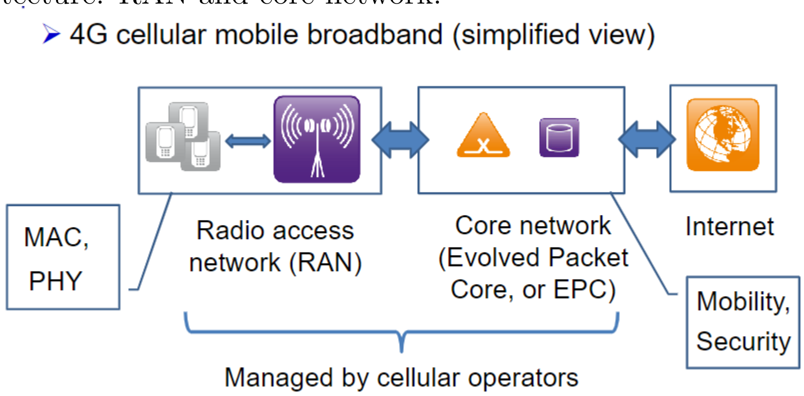
1. The modern Internet is a "network of networks". What does this mean? Draw a figure to illustrate the Internet topology as a network of different levels of operators.

   - End systems connect to Internet via access ISPs (Internet Service Providers)
   - Access ISPs in turn must be interconnected
   - Resulting network of networks is very complex



   - "tier-1" commercial ISPs (e.g., T-Mobile, AT&T), national & international coverage
   - content provider network (e.g., Google): private network that connects it data centers to Internet, often bypassing tier-1 and regional ISPs

2. Understand the function separation between the two major modules in a cellular architecture: RAN and core network.



   - Core Network manage the RAN
   - Connect RAN to Internet
   - Collection of gateway server

3. Qualitative understanding of how cellular networks' PHY layer technologies evolved from 2G to 5G.

	Bit-rate	Spectrum	bandwidth	Modulation	MAC features
2G-GSM	< 1Mbps	890-915 MHz uplink, 935-960 MHz downlink)	50 MHz	GMSK	TDMA, FDMA
3G-UMTS	~1Mbps	1.8-2.4GHz	1.25, 5 or 20 MHz	Spread spectrum	CDMA
4G	~10 Mbps	2-6GHz	Up to 100 MHz	OFDM	OFDMA
5G	>100Mbps	Including spectrum at >24 GHz (millimeter-wave spectrum)	Up to GHz	OFDM	SDM

4. What is mobile SDN and mobile edge computing? How do these technologies address the limitations of current cellular network architecture?
   Mobile software defined networking
   - Split the control and data plane for bottleneck entities
   - Centralize the control functions to ease coordination
   - Control plane should have low latency; data have high throughput
   - Virtualize gateways and other core network entities
   - eparation of concerns, specifically, separation of control and data channels for higher flexibility, better scalability, and lower cost

   Mobile edge computing
   - Move cloud servers' functions down to the "edge" servers that are close to the base stations
   - MEC servers provide computation, communication, and storage services
   - Virtualize server resources for flexibility and scalability
   - Localized or proximity processing avoids shared communication bottleneck.

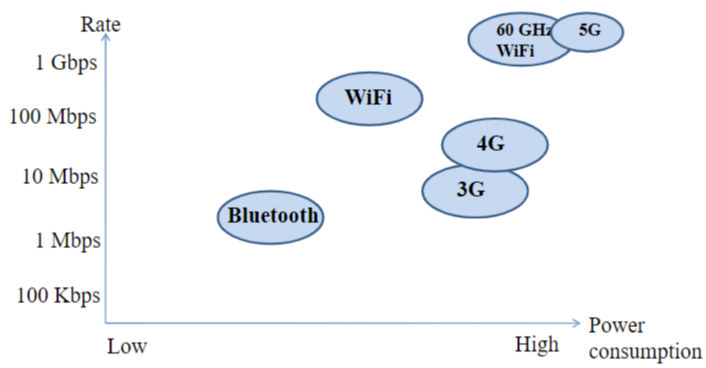
5. Qualitative understanding of how WiFi networks' PHY layer technologies evolved (802.11a/g→802.11n→802.11ac and 802.11ad).

	Bit-rate	Spectrum	Bandwidth	Modulation	Spatial streams
802.11a/g	54 Mbps	2.4 GHz and 5 GHz	20 MHz	OFDM	1
802.11n	Up to 600 Mbps	2.4 GHz and 5 GHz	Up to 40 MHz	OFDM + MIMO	Up to 4
802.11ac	Up to 6 Gbps	5 GHz	Up to 160 MHz	OFDM + MIMO	Up to 8
802.11ad	Up to 7 Gbps	60 GHz	Up to 2 GHz	OFDM + phased-array beamforming	1
802.11ax	Up to 10 Gbps	2.4, 5, 6GHz	Up to 160 MHz	OFDM + MIMO	Up to 8

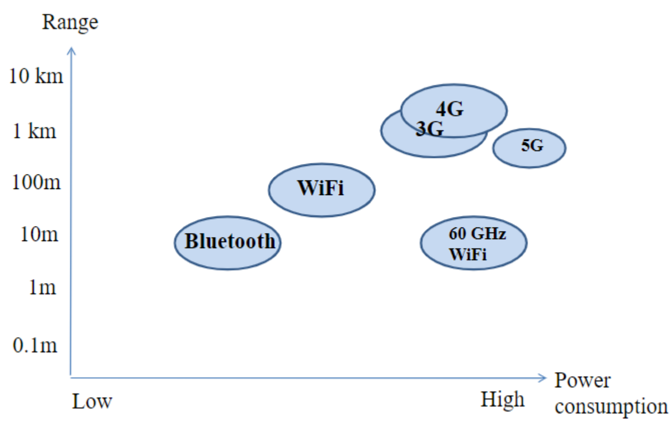
Lecture 3 Wireless Internet workflow; Mobile Internet workflow

1. Qualitative understanding of the differences between WiFi, Bluetooth, cellular networks, and 60GHz WiFi (bit-rate, range, power consumption).

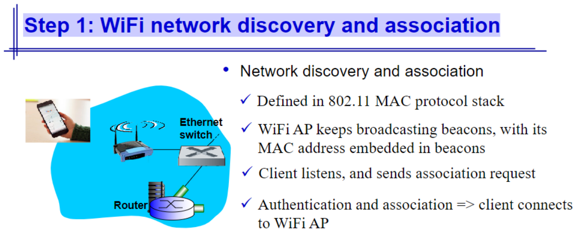
➤ Data rate vs. power consumption

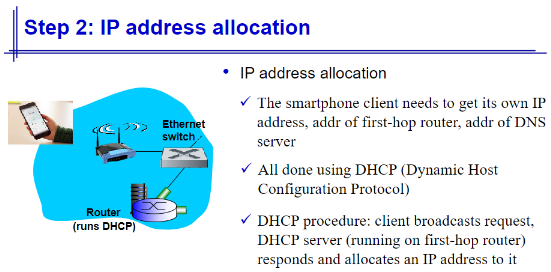


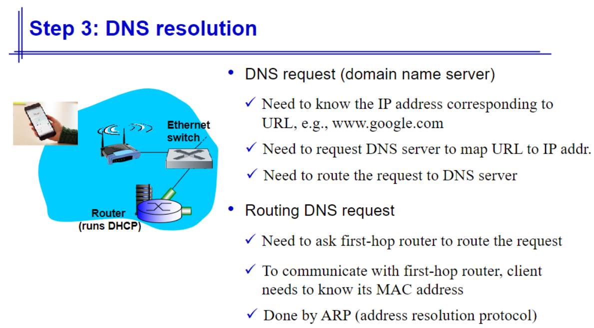
· Range vs. power consumption

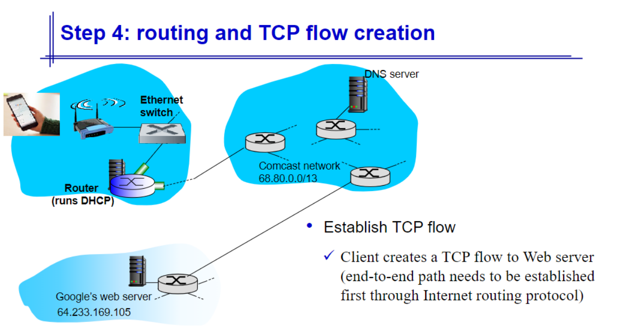


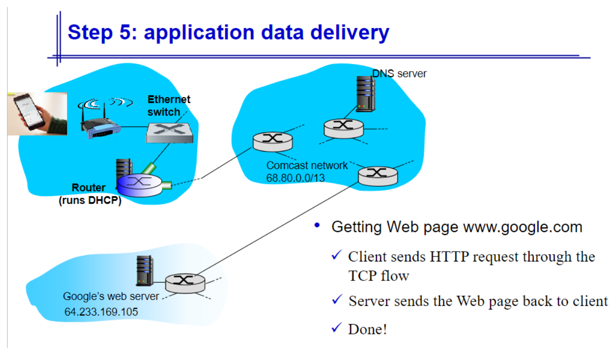
2. Describe the major protocols involved in a mobile Web browsing session, and their workflow.





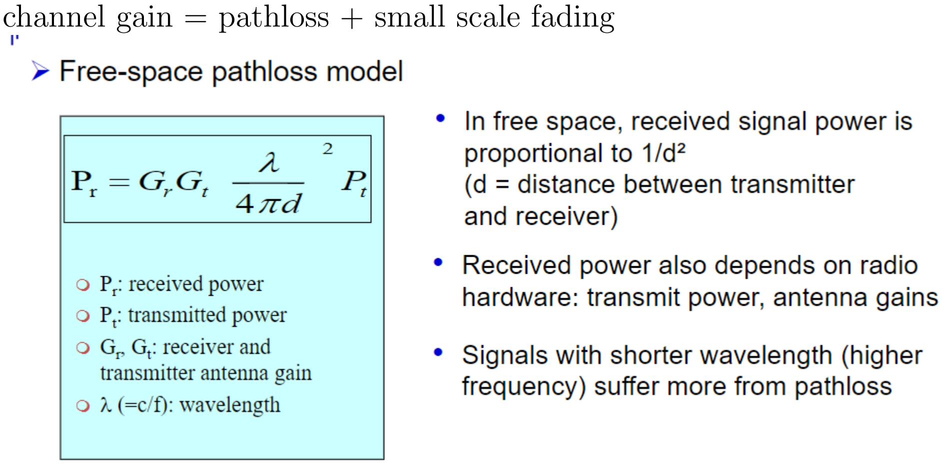






Lecture 4 Wireless channel model; digital modulation

1. Understand the freespace pathloss model.
   channel gain = pathloss + small scale fading

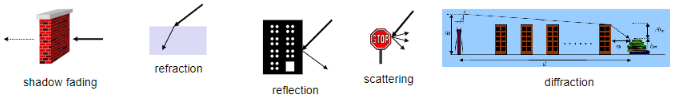


2. Understand the causes and effects of small-scale fading (Doppler fading, multipath

fading).

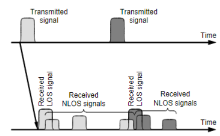
➢ Multipath fading

- shadowing (e.g., through a wall or a door);
  refraction depending on the density of a medium;
  reflection at large obstacles;
  scattering at small obstacles;
  diffraction at edges.
- All these distorted versions add up at the receiver



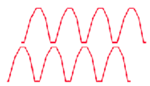
➢ Multipath fading causes inter-symbol interference

- Multiple copies of the same signal are received, each following a different path, and arriving at different times
- Distortion in time-domain: inter-symbol interference



➢ Multipath causes frequency-selective fading

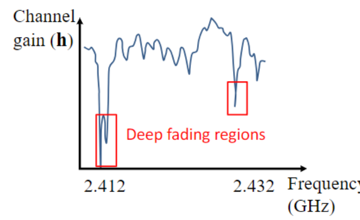
- Copies can either strengthen or weaken each other, depending on whether they are in or out of phase
- For a given receiver location, whether two waves are in or out of phase depends on frequency



- Changes of half-wavelength affect the result
  - ✓ e.g., WiFi: carrier signals at 2.4GHz, 12cm wavelength
  - ✓ This is why small adjustments in location or orientation of the wireless devices can result in big changes in signal strength

➢ Multipath effect causes frequency-selective fading

Example: channel gain of a WiFi channel



Small-scale fading: Doppler effect

➢ Doppler fading

- Movement by the transmitter, receiver, or *objects* in the environment can create a Doppler shift. Maximum doppler shift f_d occurs when the wave is traveling along the moving direction

$$f_d = (v/c) \cdot f$$

  - v is moving speed, f is center frequency of signals
- Doppler shift distorts the frequency of the signals
- Similar to change of sound frequency of a moving object (e.g., train/car passing by)
  https://en.wikipedia.org/wiki/File:Speeding-car-horn_doppler_effect_sample.ogg

3. Understand the quantitative measures of small-scale fading and their relationship

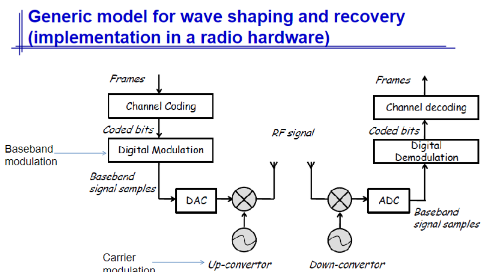
➢ Coherence time: stability of channel over time

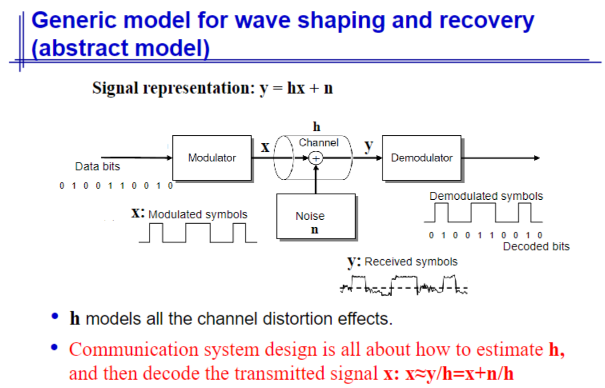
- Time duration over which the received signal experiences consistent amplitude/phase distortion
- Depends on movement speed
- Empirical model: $T_c = \frac{9}{16\pi f_d} = \frac{9c}{16\pi \cdot vf}$ $\quad f_d$ is Doppler shift

➢ Coherence bandwidth: stability of channel over frequency

- Frequency range within which two signals experience similar amplitude/phase distortion
- Depends on multipath fading effect
- Empirical model: $f_c = \frac{1}{D}$ $\quad$ D is delay spread

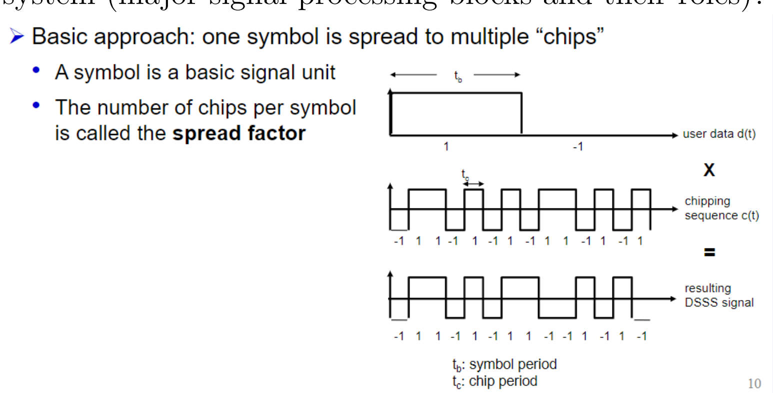
4. Understand the generic model for digital modulation (major stages and their objectives)



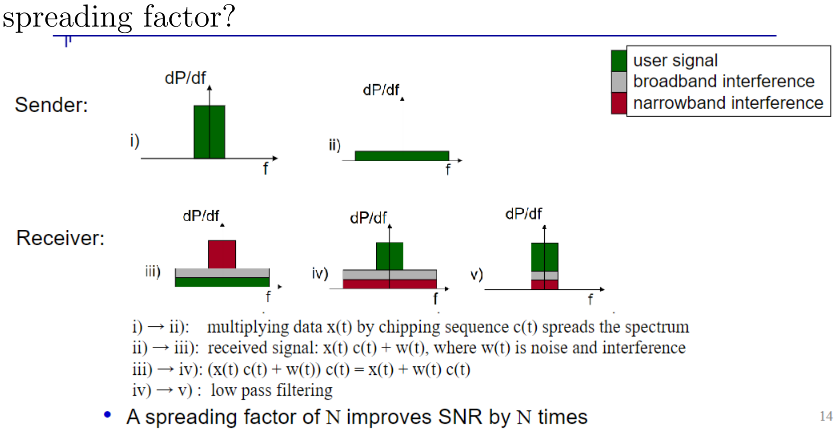


Lecture 5 Wireless channel model; digital modulation

1. How does DSSS spreading work? How is it implemented in a practical communication system (major signal processing blocks and their roles)?

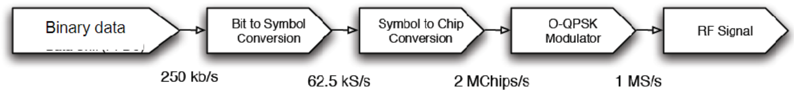


2. Why can DSSS improve SNR? What's the quantitative improvement as a function of spreading factor?

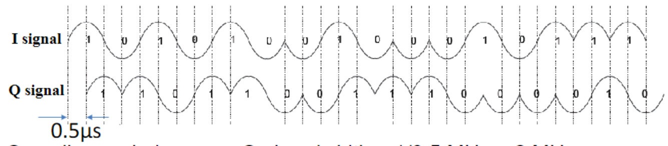


3. How is DSSS implemented in ZigBee?

- Modulation work flow



- O-QPSK modulation: digital chips to analog waves
- E.g., chip sequence 1101100111……

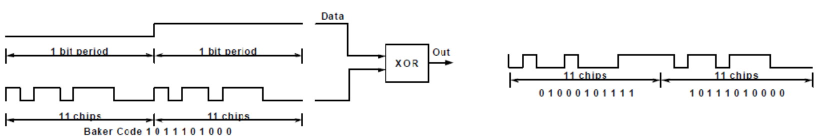


- Sampling period = 0.5µs; So bandwidth ≈ 1/0.5 MHz = 2 MHz

17

4. How is DSSS implemented in 802.11b? Why is sampling time synchronization needed and how is it achieved?

- Each bit is spread using a chip sequence (length=11) called Barker code

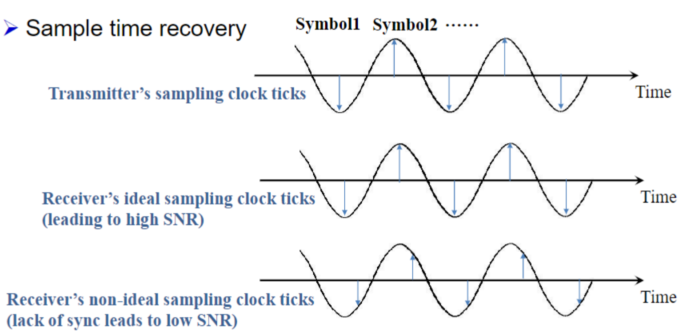


- Each chip is modulated into baseband waveforms, using DBPSK
- On the waveform, 2 samples represent 1 chip; chip rate 11MHz => Sampling rate 22 Msps, bandwidth 22 MHz; Bit-rate 1Mbps
- How to achieve higher bit-rate?
  - ✓ Use multiple orthogonal chip sequences instead of a single Barker code

19

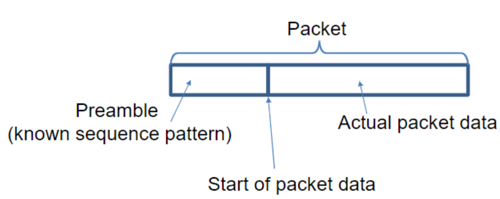
Time synchronization in 802.11b

- Sample time recovery

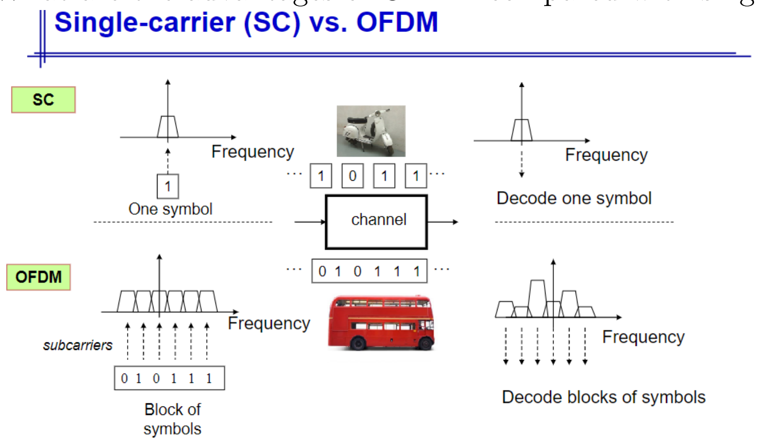


- So, sub-symbol level synchronization is needed!

- Sampling time recovery: practical implementation in 802.11b
  - Correlation with barker code, and find the peak
  - Packet detection and packet-level synchronization is also needed



5. What are the advantages of OFDM compared with single-carrier (SC) FDM?



6. Mathematical model of OFDM implementation. How is orthogonality achieved among subcarriers?

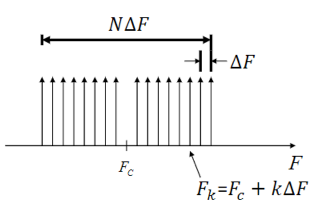
- N-point IFFT translates N frequency-domain sub-carrier symbols to a time-domain OFDM symbol (containing multiple samples, lasting T_{symbol} duration)

$$x(t) = \sum_{k=-\frac{N}{2}}^{k=\frac{N}{2}-1} X[k]\, e^{j2\pi k\Delta F t}, \quad 0 \le t \le T_{symbol}$$

$X[k]$: data symbol carried by k−th subcarrier

ΔF: subcarrier frequency offset;
$\Delta F = \frac{B}{N}$, where B is channel bandwidth
$k\Delta F$: frequency of k−th subcarrier

- Why are the subcarriers orthogonal?



Orthogonality:

$$\frac{1}{T_b}\int_{t_0}^{t_0+T_b} e^{j2\pi F_k t} e^{-j2\pi F_\ell t}\, dt = \frac{1}{T_b}\int_{t_0}^{t_0+T_b} e^{j2\pi(k-\ell)\Delta F t}\, dt = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{if } k \neq \ell \end{cases}$$

7. Why is cyclic prefix needed? How is it implemented in OFDM?

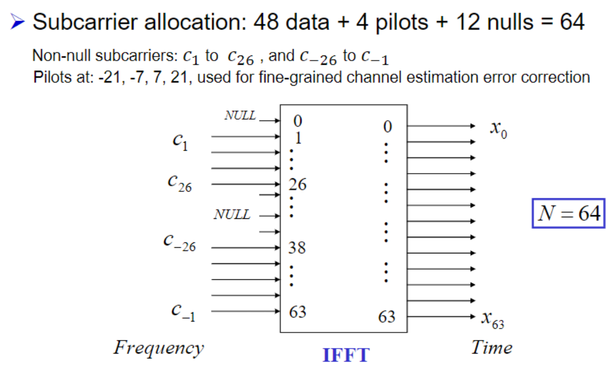
- A "guard interval" between symbols to contain inter-symbol interference
- CP is long enough to ensure the multipath from one symbol does not affect the next
- Handling time synchronization error

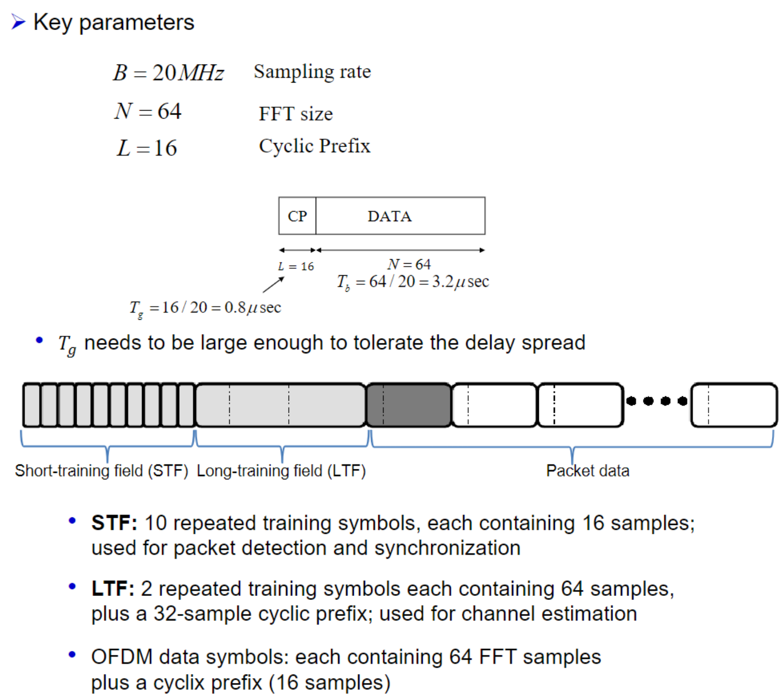
How is it implemented in OFDM?

- Copy G tailing samples to the front and form a (G+N) sample OFDM symbol

Lecture 6. WiFi packet processing

1. In 802.11 OFDM, how are bits mapped to subcarriers? Why is it done in this way? How is an OFDM symbol constructed?





Pilots at: -21, -7, 7, 21, used for fine-grained channel estimation error correction
Null subcarriers, serving as guardband to avoid leakage interference to adjacent channels

2. How is the STF preamble constructed to facilitate packet detection? How to detect the STF preamble? How is time synchronization done based on STF?
STF: 10 repeated training symbols, each containing 16 samples; used for packet detection and synchronization

➤ Packet detection

- Idea: utilizing the repeated pattern in STF
- Algorithm: self-correlation, detect if a pattern has been repeated (noise is unlikely to repeat itself)

Received digital signals:

$$y(n) = h(n)x(n) + w(n), n = 1,2 \ldots$$

Self-correlation, spanning across 16 samples (one STF symbol): $R(m) = \sum_{i=m}^{m+16-1} y(i)y^*(i-16) \approx \sum_{i=m}^{m+16-1} |y(i)|^2$

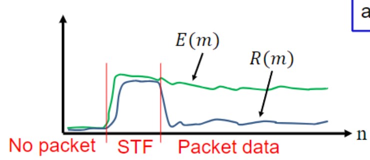
Energy of 16 samples: $E(m) = \sum_{i=m}^{m+16-1} |y(i)|^2$

➤ Packet detection

Self-correlation of non-STF signals: $R(m) = \sum_{i=m}^{m+16-1} y(i)y^*(i-16) \approx 0$

Energy of non-STF samples: $E(m) = \sum_{i=m}^{m+16-1} |y(i)|^2$

Decision logic: Packet detected if $R(m) \gg 0$ and $R(m) \approx E(m)$ and this remains true across 160-16 samples



➤ Time synchronization

- Idea: utilize the known sequence of samples in each STF symbol
- Algorithm: cross-correlation, detect if the known sequence has occurred

Received digital signals:

$$y(n) = h(n)S(n) + w(n), n = 1,2 \ldots, 16$$ (Note: S(n) is the known STF sequence)

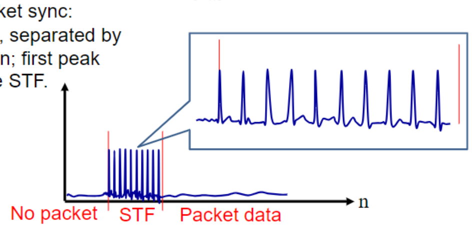
Cross-correlation, spanning across 16 samples (one STF symbol): $R(m) = \sum_{i=m}^{m+16-1} y(i)\, S^*(i-m) \approx 16\, h(m)$

Note: channel is coherent within short period, so $h(m) \approx h(m+1) \approx h(m+15) \approx \ldots$

➤ Time synchronization

Cross-correlation of non-STF samples: $R(m) = \sum_{i=m}^{m+16-1} y(i)\, S^*(i-m) \approx 0$

Decision logic for packet sync: 10 consecutive peaks, separated by 16 samples in between; first peak point is the start of the STF.



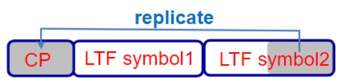
3. How is the LTF preamble constructed? How to use the LTF to estimate frequency offset?
   LTF: 2 repeated training symbols each containing 64 samples,plus a 32-sample cyclic prefix; used for channel estimation

- Frequency offset estimation
  - Utilize the LTF, which has two repeating blocks, each N=64 samples

$$y(n+N) = y(n)e^{j2\pi N\Delta f}$$

$$\text{So,} \quad e^{j2\pi N\Delta f} = \frac{y(n+N)}{y(n)} \quad \text{and} \quad \Delta f = \mathrm{Img}\left\{\frac{y(n+N)}{y(n)}\right\}/(2\pi N)$$

  - We can compute a Δf for each sample in an LTF symbol, and then take the average



- Frequency offset compensation
  - i-th sample needs to multiply $e^{-j2\pi i\Delta f}$

4. How is channel equalization done for 802.11 OFDM packets?

- Channel gain estimation
  - Assume flat fading (each subcarrier k is distorted by a single multiplier $H(k)$, which is the channel gain)
  - Use LTF to estimate $H(k)$

    For LTF, received signal on subcarrier k (frequency domain) is:

    $Y(k) = FFT(LTF\ symbol\ 1)$ or $FFT(LTF\ symbol\ 2)$

    The signal sent on subcarrier k is known, denoted as $Sl(k)$

    Then the channel gain on subcarrier k is: $H(k) = Y(k)/Sl(k)$
- Channel equalization for subsequent OFDM data symbols in the packet:

  To decode the data symbol carried on subcarrier k, simply do: $X(k) = Y(k)/H(k)$

Lecture 7. MIMO

1. Understand the fundamental differences between diversity gain (from SIMO or MISO) and multiplexing gain (from MIMO).

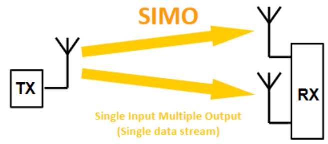
   Diversity gain(Receiver Diveristy and Transmit diversity):

   - Focus: Mitigating fading, improving link reliability.
   - Configuration: SIMO or MISO.
   - Spatial Separation: Different channels experienced by multiple antennas.
   - Objective: Enhancing robustness against channel variations

   Multiplexing gain:

   - Focus: Increasing data rate through simultaneous transmission of multiple data streams.
   - Configuration: MIMO.
   - Spatial Separation: Exploiting spatial diversity for independent communication paths.
   - Objective: Maximizing data throughput by using spatial resources efficiently.

2. How does SIMO improve wireless link SNR?



- Receiver coherently combines signals received by multiple antennas
- Asymptotic gain: Increasing SNR proportionally to Nr (#of receiving antennas)
  - Intuition: received signal power adds up
- What's the capacity gain, given the SNR improvement?
  - Can be analyzed based on Shannon's equation: $C = B\log_2(1+SNR)$
  - When SNR is low, $\log_2(1+SNR) \approx SNR$, so gain is almost linear w.r.t. Nr
  - When SNR is high, gain is approximately log w.r.t. Nr

3. How does open-loop transmit diversity improve SNR?
Send redundant versions of the same signal (symbol), over multiple time slots, and through multiple antennas
Encode the symbols differently for different time slots and TX antennas

Open-loop transmit diversity

- Example: STBC with 2 TX antenna

$$\boxed{\begin{aligned} r(t_1) &= h_1 s_1 + h_2 s_2 \\ r(t_2) &= -h_1 s_2^* + h_2 s_1^* \end{aligned}}$$

  - Receiver side: diversity combining

$$y_1 = h_1^* r(t_1) + h_2 r^*(t_2) = (|h_1|^2 + |h_2|^2)s_1$$
$$y_2 = h_2^* r(t_1) - h_1 r^*(t_2) = (|h_1|^2 + |h_2|^2)s_2$$

  - More generally, with N_t TX antennas, combined SNR at receiver:

$$\gamma_\Sigma = \frac{\mathcal{E}_x}{\sigma^2}\frac{|h_1|^2 + |h_2|^2 + \ldots + |h_{N_t}|^2}{N_t} \rightarrow \frac{\mathcal{E}_x}{\sigma^2} E[|h_i|^2]$$

  (Note: total TX power, $\mathcal{E}_x$, is split among N_t transmit antennas)

  - Open-loop transmit diversity causes the received SNR to "harden" to the average SNR. In other words, it eliminates the effects of small-scale fading but does not increase the average received SNR.

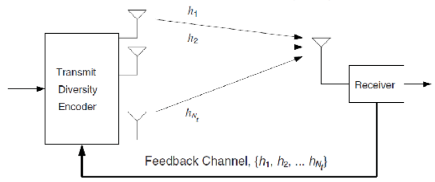
4. How does close-loop transmit diversity improve SNR and link capacity?
Send redundant versions of the same signal (symbol), over the same time slot
Encode the symbols differently for different TX antennas

Closed-loop transmit diversity

- Asymptotic gain from closed-loop transmit diversity
  - ✓ Suppose we have 2 transmit antennas, then instead of receiving symbol x, we receive: x+x=2x, received power becomes $4|x|^2$,
    SNR increases to 4 times! ?
    But under the practical constraint that total transmit power remains constant, then we receive $\frac{x}{\sqrt{2}} + \frac{x}{\sqrt{2}} = \sqrt{2}\,x$. Received power becomes $2|x|^2$.
    SNR actually doubles in practice.
  - ✓ More generally, with N_t TX antennas, SNR increases to N_t times, assuming total transmit power remains constant (splits among N_t antennas)
  - ✓ What's the capacity gain?

➢ Working principle

- Send redundant versions of the same signal (symbol), over the same time slot
- Encode the symbols differently for different TX antennas
  - ✓ i.e., weight the symbols on different antennas, following a precoding algorithm
  - ✓ Precoding design requires feedback of channel state information (CSI)



Lecture 8. MU-MIMO and network MIMO

1. How does MIMO spatial multiplex gain increase with the number of transmit or receive antennas?
   In general, capacity gain from spatial multiplexing scales linearly with $Min(N_t, N_r)$. N_t is number of variables and N_r is numbers of eqns

2. How does channel rank and channel condition number affect MIMO spatial multiplexing gain? Understand the high-level mathematical reasons behind.

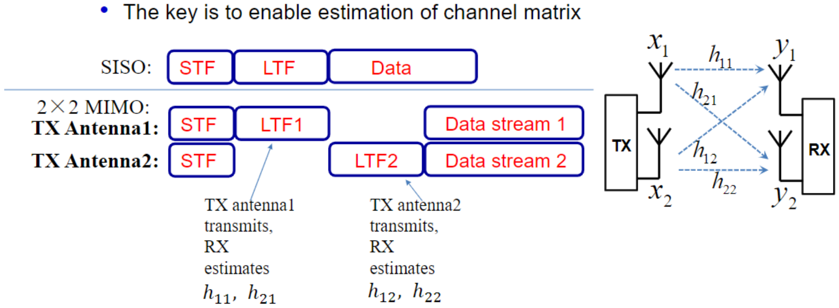
   ➢ In practice

   - Spatial multiplexing gain also depends on channel "condition"
     - ✓ If the channels between different antennas are correlated, e.g., are all the same, then you can't solve the equations. Spatial multiplexing becomes infeasible!
     - ✓ Intuition: the equations are correlated, then the "system of equations" is unsolvable
   - Practical wireless devices' multiple antennas are separated sufficiently far (further than half-wavelength), in order to minimize the correlation

3. How does 802.11 MIMO estimate the channel matrix?

   ➢ SU-MIMO packet construction and channel estimation

   - The key is to enable estimation of channel matrix



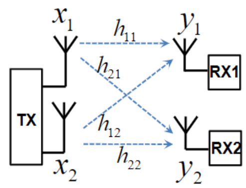
4. How does MU-MIMO enable concurrent transmission from a single AP to multiple users?

- Use TX precoding to remove cross-talk
  - Send a weighted mix of x_1 and x_2

    TX antenna1 sends: $w_{11}x_1 + w_{12}x_2$

    TX antenna2 sends: $w_{21}x_1 + w_{22}x_2$
  - Data received by RX1:

    $$y_1 = h_{11}(w_{11}x_1 + w_{12}x_2) + h_{12}(w_{21}x_1 + w_{22}x_2)$$
    $$= (h_{11}w_{11} + h_{12}w_{21})x_1 + (h_{11}w_{12} + h_{12}w_{22})x_2$$



  - RX1 only wants x1, so ideally, we should force $(h_{11}w_{12} + h_{12}w_{22}) = 0$

- MU-MIMO precoding
  - TX can obtain h_{11}, h_{12} from RXs' feedback, so it can tune w_{12}, w_{22} to satisfy $(h_{11}w_{12} + h_{12}w_{22}) = 0$
    - This cancels the cross-talk interference from x2 to x1
    - Similarly, we can cancel that from x1 to x2.

      Data received by RX2: $y_2 = h_{21}(w_{11}x_1 + w_{12}x_2) + h_{22}(w_{21}x_1 + w_{22}x_2)$
      $= (h_{21}w_{11} + h_{22}w_{21})x_1 + (h_{12}w_{12} + h_{22}w_{22})x_2$

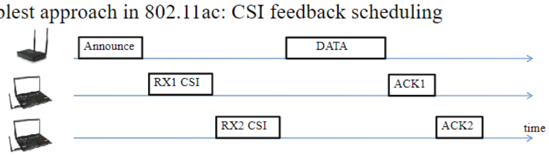
      Nullify x1: $(h_{12}w_{11} + h_{22}w_{21}) = 0$
    - Essentially, the TX precoding weights w_{**} need to meet two zero-forcing equalities:

      $(h_{11}w_{12} + h_{12}w_{22}) = 0$ $(h_{12}w_{11} + h_{22}w_{21}) = 0$

      This is called Zero-Forcing Beamforming (ZFBF)

How does TX obtain channel state information (CSI), i.e., h_{**}

✓ Simplest approach in 802.11ac: CSI feedback scheduling



5. What's the asymptotic capacity gain of MU-MIMO, as a function of number of antennas/users?

- Asymptotic capacity gain
  - If the transmitter has N_t antennas, then it can send N_t streams of data simultaneously to N_t users, increasing capacity to N_t times compared with single-antenna transmitter

- Limitation
  - MU-MIMO is essentially a form of spatial multiplexing
  - So the channel must be well-conditioned: multipath rich, and full-rank channel matrix
  - MU-MIMO requires closed-loop feedback---non-trivial overhead, making the net throughput gain lower than N_t times

6. Where does the MU-MIMO overhead come from? How does it affect the effective throughput?

- Overhead comes from multiple dimensions
  - Time-domain: CSI needs to be updated over time
  - Frequency-domain: CSI needs to be updated for different subcarriers
  - Actual CSI values: need to be represented in digital bits and carried through feedback packets
- Taming the feedback overhead through compression
  - Time-domain: update CSI intermittently
  - Frequency-domain: feedback CSI across a few subcarriers, and then interpolate the rest
  - Quantization: quantize the channel gain values using fewer bits

What's the catch?
More aggressive compression → less accurate CSI → lower MU-MIMO gain

14

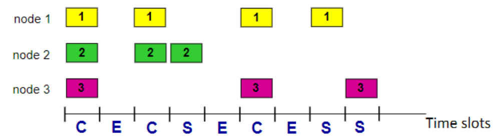
7. Understand the differences between the 3 forms of network MIMO implementation in LTE.

- Coordinated Scheduling (CoMP-CS)
  - Allocate different time and frequency resources to cell edge users served by different cells
  - E.g. in synchronized network, neighboring access nodes are using different resource blocks for serving cell edge users
- Coordinated Beamforming (CoMP-CB)
  - Allocate different spatial resources to users at cell edge, but time and frequency resources are reused
  - Beamforming weights can be calculated in a way that cell edge users served by neighboring cells are nulled (allocated zero power)
    - Requires CSI from users to be nulled
- Joint Transmission (CoMP-JT)
  - Central precoding needed for multiple access nodes (or remote radio heads)
  - Data transmitted simultaneously from multiple access nodes or RRH
  - High demand onto the backhaul since data has to be in several places
  - Tight synchronization of access nodes (frequency&phase) needed
  - Theoretically, JT can get rid of all inter-cell interference, and scale network capacity with network density!

Lecture 9. MAC overview; CSMA

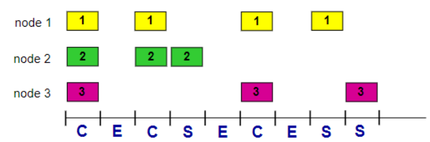
1. Slotted ALOHA and unslotted ALOHA: operations, pros and cons, and efficiency
   Slotted ALOHA

Slotted ALOHA



- Assumptions
  - Time divided into equal size slots (time to transmit 1 frame)
  - Nodes are synchronized
  - All nodes can detect collision
- Operation
  - When node obtains fresh frame, transmits in next slot
  - If no collision: send new frame in next slot
  - If collision: node retransmits frame in each subsequent slot with prob. p until success

Slotted ALOHA



- Pros
  - Single active node can continuously transmit at full rate of channel
  - Simple; highly decentralized: only need to sync each slot
- Cons
  - collisions, wasting slots; inefficient collision resolution
  - Idle slots, underutilizing channel
  - Require clock sync

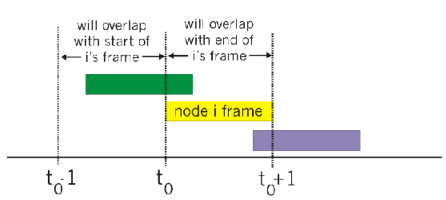
Efficiency of slotted ALOHA

- Definition of MAC efficiency: long-run fraction of successful slots (assuming many nodes, all with many frames to send)
- Suppose: N nodes with many frames to send, each transmits in slot with probability p
- Prob. that given node has success in a slot = $p(1-p)^{N-1}$
- Prob. that any node has a success: $\mathrm{E(p)} = Np(1-p)^{N-1}$

At best: channel used for useful transmissions 37 of time!

Unslotted ALOHA

- Unslotted Aloha: simpler, no synchronization (each node runs its own clock)
- When frame first arrives, transmit immediately
- Upon collision, retransmit with probability p
- Collision: frame sent at t0 collides with other frames sent in [t0-1,t0+1]



1

Pure (unslotted) ALOHA

P(success by given node) = P(given node transmits) ·
P(no other node transmits in $[t_0-1,t_0]$) ·
P(no node transmits in $[t_0,t_0+1]$)
$= p \cdot (1-p)^{N-1} \cdot (1-p)^{N-1}$
$= p \cdot (1-p)^{2N-2}$

… choosing optimum p and let N go to infinity, then efficiency
$= 1/(2e) = 0.18$

Efficiency is 50% lower compared with slotted Aloha!

2. How does collision detection and exponential backoff work in CSMA/CD? Why is exponential backoff needed?
CSMA: listen before transmit

- If channel sensed idle: transmit entire frame
- If channel sensed busy, defer transmission

Collision detection: measure signal strengths, compare transmitted vs. received signals

1. Ethernet card receives data from upper layer (running on end-host)
2. Before TX, listen:
   If: channel idle, then start TX;
   Else: wait until idle
3. During TX:
   If: detect no other TX => no collision, done!
   Else: abort
4. After abort, enter **Binary Exponential Backoff (BEB):**
   after *m*th collision, transmitter chooses *K* at random from $\{0,1,2,\ldots,2^m-1\}$.
   transmitter waits K·512 bit times, returns to Step 2
   (more collision=>longer backoff)

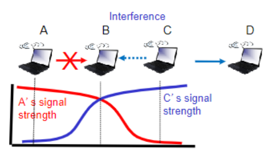
Lecture 10. CSMA

1. Pros and cons of major MAC paradigms: channel partitioning, taking turns, random access
TDMA:
Pros: Low collision rate; full channel utilization in high traffic demand.
Cons: Low channel utilization in low traffic demand; need clock synchronization.
FDMA:
Pros: Full channel utilization in high traffic demand; no clock synchronization.
Cons: Low channel utilization in low traffic demand.
Polling:
Pros: Low collision rate; easy to implement.
Cons: Overhead associated with polling; single point of failure; latency in waiting for

the central coordinator to finish polling others.
Token rings:
Pros: Low collision rate; easy to implement.
Cons: Overhead associated with passing the token; The token becomes a single point of failure; Latency in waiting for the token to become available.
Slotted ALOHA:
Pros: Easy to implement; only have to worry about a single parameter.
Cons: Collisions can occur and only 37 maximum efficiency; need clock synchronization. achievable throughput of the network?
Pure ALOHA:
Pros: Simpler than slotted ALOHA; no clock synchronization.
Cons: Collisions can occur and only 18 maximum efficiency.
CSMA/CD:
Pros: Ability to detect collision and abort transmission; decentralized implementation; no clock synchronization;scales with size.
Cons: Collisions can still occur.

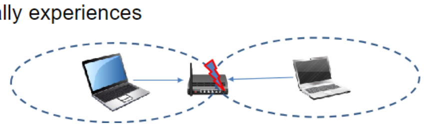
2. Why carrier sensing alone doesn't prevent collision in wireless networks? Why CSMA/CD doesn't work for wireless networks?

- Wireless channel is shared medium
  - Collision: when multiple nodes (AP/user) transmit concurrently

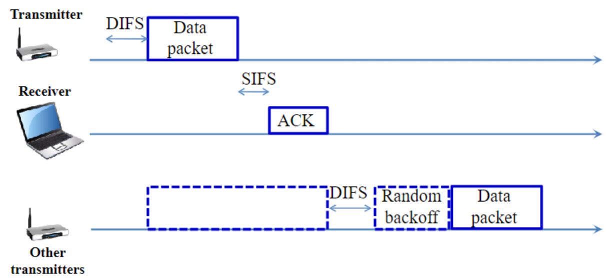


- Carrier sensing (listen before talk) is necessary
  - Node doesn't transmit if channel is busy (occupied by others)
  - But this doesn't completely remove collision
    - ✓ When two transmitters assume the channel idle at the same time

- Can we do CSMA/CD? No!
  - Wireless communication devices cannot listen while transmitting
    - ✓ Wireless nodes are usually half-duplex
    - ✓ Ratio of transmitted to received signal power typically > 60dB
    - ✓ Transmission drowns out the ability of the radio to hear a collision
  - Even if the transmitter is full-duplex (i.e., it can listen while transmitting), what it hears may be different from what the receiver actually experiences



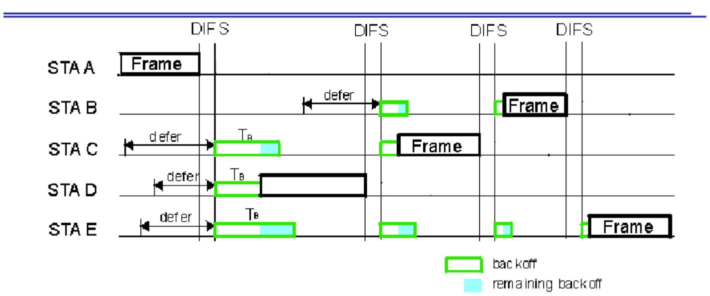
3. Understand the CSMA/CA operations in 802.11



- Forcing transmitters to wait for random amount of time can reduce the chance of collision (similar to Ethernet)
- If medium is sensed busy
  - Pick backoff counter: a random number between 0 and contention window (CW) size
  - Decrement backoff timer in each slot, until reaching 0, and then transmit
  - Freeze backoff if medium becomes busy (Why?)
  - Collision occurs if two transmitters pick the same backoff counter
- After every failed transmission attempt (assuming collision occurred)
  - Increase the CW exponentially, starting with 32 up to 1024
  - Reset CW to 32 (minimum window) upon success

4. How does random backoff work in 802.11 CSMA/CA?

Random backoff in 802.11

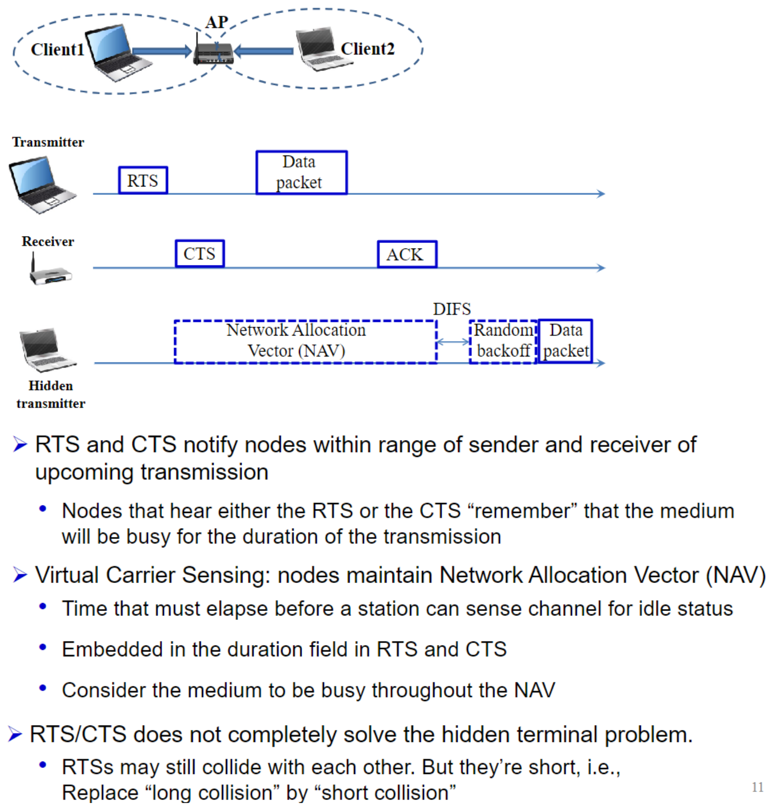


- Stations (STAs) listen to the channel before transmission after DIFS
- Avoid collision by random backoff
- Start transmission when backoff timer fires; Freeze backoff timer when others start transmission; continue backoff when channel becomes idle;

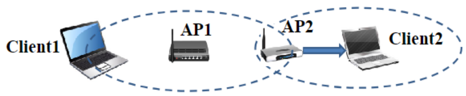
1. Each STA maintains a contention window (CW)
   - Initialized to CW_{min} = 32
2. Randomly pick a number, say *k*, between [0,CW-1]
3. Count down from *k* when the channel becomes idle
4. Freeze backoff countdown if the channel becomes busy during backoff
5. Start transmission when *k = 0* if the channel is still idle
6. Double CW for every unsuccessful transmission, up to CW_{max} (1024)
7. reset CW to CW_{min} after every successful transmission

5. What's hidden terminal? How does it affect WiFi performance? How does RTS/CTS remedy the problem?

Considering a scenario shown in Fig. 2, two stations (clients) who are out of range of each other but the common recipient (AP) is inside the range of the two station. Since the stations cannot hear each other, the two stations may transmit at the same time since they think the channle is idle. Simultaneous transmission will cause collision at the common recipient which leads to hidden terminal problem.
RTS/CTS does not fully solve the problem since RTSs may still collide with each other.
How RTS/CTS Work
Sender first transmits small request-to-send (RTS) packets using CSMA
AP broadcasts clear-to-send (CTS) in response to RTS
CTS heard by all nodes (AP can reach everyone)
Sender transmits data frame
Other stations defer transmissions



6. What’s exposed terminal? How does it affect WiFi performance?
Considering a scenario shown in Fig. 3, the exposed terminal problem occurs when two transmitters (APs) intending to transmit data to two different receivers (clients) can detect transmission of each other’s signals. Therefore, when AP2 transmit signal to client2, AP1 will backoff since it detects the transmission from AP2, even though their transmissions should not affect their corresponding receiver. This will reduce the spatial reuse opportunities which in turn reduces effective total network throughput.
No. The RTS/CTS mechanism will force the exposed devices to hold transmission. There are a few methods to tackle this issue through spatial reuse. If you are interested, please check the spatial user in 802.11ax and NOMA



7. Understand the differences between bit-rate, capacity, and throughput.
   Bit rate: Bit-rate, also known as data rate or transmission speed, refers to the number of bits transmitted per unit of time. It is a measure of how quickly data is transmitted over a communication channel.
   Capacity:Capacity refers to the maximum amount of data that a communication channel or network link can handle. It represents the upper limit of data transfer capability under ideal conditions.
   Throughput:Throughput refers to the actual amount of data that is successfully transmitted over a communication channel or network link in a given period. It reflects the real-world performance, which may be less than the theoretical capacity due to various factors such as network congestion, protocol overhead, and errors.

8. What's frame fragmentation and aggregation? How do they affect 802.11 network performance?
   - Success probability vs. frame size
     - Large frame reduces overhead, but is less reliable
     - Discard the frame even if only one bit is in error
     - Packet delivery ratio of an N-bit packet: $(1\text{-BER})^N$
   - Fragmentation
     - Break a frame into small pieces
     - A burst of bit-errors only affects small fragments
   - Aggregation
     - Aggregate multiple small frames in order to reduce the overhead
     - But minor error may corrupt the whole big frame

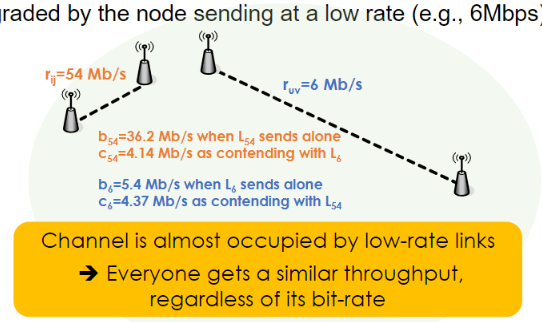
9. Qualitative understand of the relation between SNR and link throughput.

10. What are the major challenges for rate adaptation in 802.11?
    - Channel quality changes very quickly
      - Especially when the device is moving
    - Can't tell the difference between
      - poor channel quality due to interference/collision (high noise level)
      - poor channel quality due to distance (low signal level)

    Ideally, we want to decrease the rate due to low signal strength, but not interference/collision.

11. Cause and effect of the rate anomaly in 802.11

- The throughput of a node sending at a high rate (e.g., 54Mbps) is degraded by the node sending at a low rate (e.g., 6Mbps)



- Root causes
  - 802.11 supports multiple transmission bit-rates, each of which has a different modulation and coding scheme
  - 802.11 ensures packet fairness, instead of time fairness

> Packet fairness: each STA has an equal probability to win the contention → the average number of delivered packets for all nodes are roughly the same (802.11)
>
> Time fairness: each node occupies roughly the same proportion of channel time

(B is the packet size in bytes)

Lecture 11. MAC models

1. Basic concept of stochastic process, Markov process, Markov chain.

   - Random process
     - Random process: a function whose output varies over time and is defined by a set of random variables
   - Markov process
     - A memoryless random process -- the value of the random variable at instant n depends only on its immediate past value at instant n -1
     - The value of random variable S(n) represents the state of the system at a given time instance n
   - Markov chain
     - A Markov process with a discrete state space
     - Memoryless property still holds: S(n) is a function of S(n-1) only

2. Markov chain concepts: transition probability, transition matrix, equilibrium state.

   - Transition probability $p_{ij}(n)$: probability that system is in state i at time n, given that it was in state j at past time n-1, i.e.,

     $$p_{ij}(n) = p[S(n) = i \mid S(n-1) = j]$$

   - If transition probability is independent of time, i.e.,

     $$p_{ij} = p[S(n) = i \mid S(n-1) = j]$$

     then the Markov chain is a homogeneous Markov chain.

- Denote $s_i(n)$ as the probability that the system lies in state i in any time instance n, i.e., $s_i(n) = P[S(n) = i]$. Also, suppose there are m states in total. Then,

$$s_i(n) = \sum_{j=1}^{m} p_{ij}\, s_j(n-1)$$

- We can express this in matrix form as:

$$\boldsymbol{s}(n) = \boldsymbol{P}\, \boldsymbol{s}(n-1)$$

where $\boldsymbol{P} = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{bmatrix}$ is the state transition matrix.

$\boldsymbol{s}(n) = [s_1(n) \;\; s_2(n) \ldots \;\; s_m(n)]^T$ is the state vector.

23

- At steady state, as $n \to \infty$, the transition probabilities no longer depends on n, and the state vector settles down to a fixed value, which satisfies:

$$\boldsymbol{P}\, \boldsymbol{s} = \boldsymbol{s}$$

- This also implies that:

The probability that the system lies in a certain state, $s_i(n)$, no longer depends on time n

The components in $\boldsymbol{P}^n$, no longer depend on time n.

Each row in $\boldsymbol{P}^n$ comprises identical values.

- The state of the system at equilibrium or steady state can then be used to obtain performance parameters such as throughput, delay, etc.

3. Basic properties of state vector, transition matrix.

- Some properties
  - Because state vector $\boldsymbol{s}(n)$ describes the probabilities of all possible states, we have:

$$\sum_{i=1}^{m} s_i(n) = 1, n = 0,1, \ldots$$

  - The column of $\boldsymbol{P}$ represents transition out of a given state, so we have:

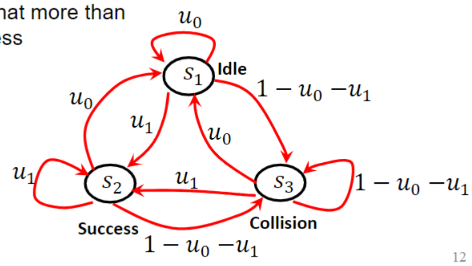
$$\sum_{i=1}^{m} p_{ij} = 1$$

i.e., the column covers all possible transition events out of that state j

4. Construct a Markov chain transition matrix for a given Markovian problem
Refer to Homework3 Solution

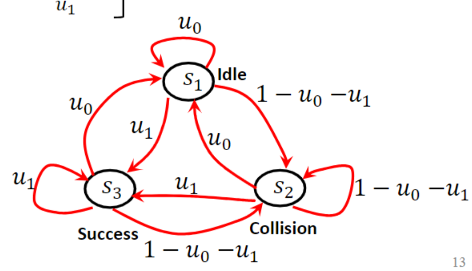
5. Model slotted ALOHA as a Markov chain and solve for its steady state; model slotted ALOHA performance based on the steady-state model

- Notations
  - u_0: Probability that all users are idle.
  - u_1: Probability that exactly one user attempts to access.
  - $1 - u_0 - u_1$: Probability that more than one users attempt to access



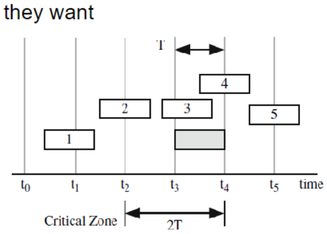
- Transition matrix

$$P = \begin{bmatrix} u_0 & u_0 & u_0 \\ 1-u_0-u_1 & 1-u_0-u_1 & 1-u_0-u_1 \\ u_1 & u_1 & u_1 \end{bmatrix}$$

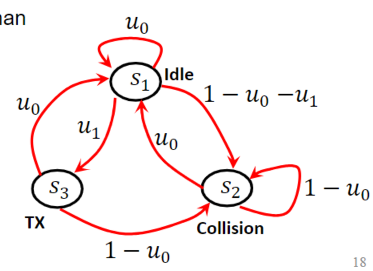


6. Model unslotted ALOHA using Markov chain

- Assumptions
  - users are allowed to transmit at any time they want
  - All frames have equal duration T
  - Transmission succeeds if a frame is sent at time t and there are no other transmissions starting during the time period t-T to t + T.
  - So for a successful transmission at a given time step, the channel must be quiet and all other users must be idle for 2T.



- Notations
  - u_0: Probability that all users are idle.
  - u_1: Probability that exactly one user attempts to access.
  - $1 - u_0 - u_1$: Probability that more than one users attempt to access



➢ Transition matrix

$$P = \begin{bmatrix} u_0 & u_0 & u_0 \\ 1-u_0-u_1 & 1-u_0 & 1-u_0 \\ u_1 & 0 & 0 \end{bmatrix}$$

7. Model CSMA/CD and CSMA/CA (under certain simplifications) using Markov chains

CSMA/CD:

Modeling CSMA/CD

➢ Assumptions and notations

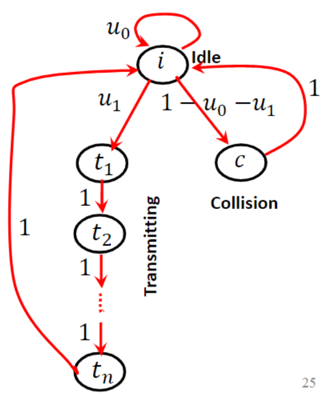
- N users; frame duration τ_t
- Time step of Markov chain equals the collision detection latency, i.e., T= τ_c. Also, n = $\tau_t/\tau_c \gg 1$
- Probability that a user attempts transmission equals a (a<1, so this is implicitly modeling a random backoff strategy—Longer backoff window is equivalent to smaller a)
- A user can have at most one frame waiting for transmission (i.e., no queueing). Frame arrival probability equals a
- Simplification: All users have the same contention window size (instead of using binary exponential backoff)

➢ Notations

- u_0: Probability that all users are idle.
- u_1: Probability that exactly one user attempts to transmit.

$$u_k = \binom{N}{k} a^k (1-a)^{N-k}$$

- t_n: Once started, a transmission takes n slots to complete



25

➢ Transition matrix

$$P = \begin{bmatrix} u_0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 \\ u_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 1-u_0-u_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

➢ State vector

$$\mathbf{s} = [\, s_i \quad s_{t_1} \quad s_{t_2} \quad \ldots \quad s_{t_n} \quad s_c]^T$$

CSMA/CA:

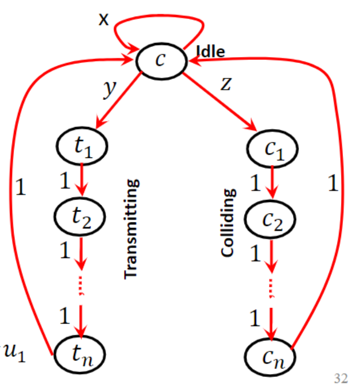
- Assumptions and notations
  - N users; frame duration τ_t; Maximum backoff window size is ω slots
  - Duration of one contention window slot equals the propagation delay plus carrier sensing delay (i.e., time to sense the channel to be busy/idle)
  - Time step of Markov chain equals one contention window slot
  - Duration of one frame equals n contention window slots
  - Probability that a user has a frame waiting to be transmitted at any time slot equals a
  - A user can have at most one frame waiting (i.e., no queueing)
  - Simplification: All users have the same contention window size ω (instead of using binary exponential backoff)

30

- Assumptions and notations
  - x: probability that no users have frame to send across the entire contention window (w slots)

    $x = u_0^w$
  - y: probability that only one user in any of the reservation slots requests a transmission and all users in the previous slots did not attempt to access the channel.

    $y = u_1 + u_0 u_1 + u_0^2 u_1 + \cdots + u_0^{w-1} u_1$
    $= u_1(1 - u_0^w)/(1 - u_0)$



32

- Transition matrix

$$P = \begin{bmatrix} x & 0 & 0 & 1 & 0 & 0 & 1 \\ y & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ z & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

- State vector

$$\boldsymbol{s} = [\, s_i \quad s_{t_1} \quad s_{t_2} \quad \dots \quad s_{t_n} \quad s_{c_1} \quad s_{c_2} \quad \dots \quad s_{c_n}]^T$$

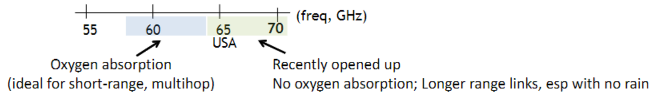
8. Model similar MAC protocols using Markov chains.

Lecture 12. mmWave networking

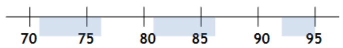
1. What's mmWave? What are the typical mmWave bands used in practical communication systems?
   Millimeter wave (mmWave) refers to a band of radio frequencies in the electromagnetic spectrum that falls within the wavelength range of approximately 1 millimeter to 10 millimeters. In terms of frequency, mmWave spans from 30 gigahertz (GHz) to 300 gigahertz (GHz).
   Shorter wavelengths, higher attenuation; Use highly directional, electronically steerable phased-arrays to overcome propagation loss

Millimeter-wave ends the spectrum crunch

- 20-30 GHz licensed mmWave band: 28 GHz, 39 GHz, etc., used by 5G NR
- 60 GHz mmWave band: 14 GHz unlicensed spectrum, available for free



- Above 70 GHz, light-licensed mmWave band: ~13 GHz of spectrum



- Spectrum beyond 100 GHz will be usable as RF hardware matures (6G?)

4

2. Understand the factors that determine the mmWave link budget

- Given a desired receiver sensitivity (i.e., received power), what is the required transmit power to attain a desired range?
  OR what is the attainable range for a given transmit power?
- Can be estimated based on free-space pathloss model:

$$P_{RX} = P_{TX}\, G_{TX} G_{RX} \frac{\lambda^2}{(4\pi d)^2} > P_{RX}(min)$$

$$P_{RX,dBm} = P_{TX,dBm} + G_{TX} + G_{RX} - L_{pathloss,dB}(d) > P_{RX,dBm}(min)$$

Note: $P_{dBm}=10\log_{10} P_{mW}$ $\quad L_{pathloss,dB}(d) = 10log_{10}\frac{16\pi^2 d^2}{\lambda^2}$

- **Receiver sensitivity** $P_{RX,\mathrm{dBm}}(\mathrm{min})$: minimum received power required to attain a desired bit-rate

$$P_{RX,dBm} = P_{TX,dBm} + G_{TX} + G_{RX} - L_{pathloss,dB}(d) > P_{RX,dBm}(min)$$

3. How does rate adaptation work in 802.11? Under the same SNR, how will the packet

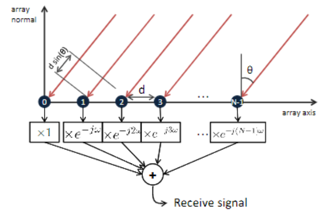
- Simple rate adaptation algorithm
  - Monitor packet error rate over a window (K packets or T seconds)
  - If loss rate > $thresh_{high}$ (or SNR < $thresh_{low}$, etc.)
    – Reduce Tx rate
  - If loss rate < $thresh_{low}$
    – Increase Tx rate

loss rate or bit error rate differ among different MCS levels?

4. How does phased-array beamforming work? What's phased-array gain pattern? How is the gain pattern affected by the phased-array size?
Large N(phased-array size) narrower beam

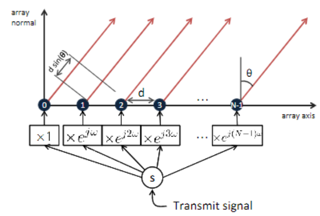
- Phased array beamforming basics
  - Receive beamforming: multiply by conjugate of phase offsets for constructive reception of signal coming from angle θ

  (ensure such signals collected by each antenna elements have 0 phase offset)



(Note: $\omega = \pi \sin\theta$)

- Phased array beamforming basics
  - Transmit beamforming: excite elements with proper phase offset to generate strong signal power in direction θ (ensure signals along direction θ coherently combine with each other, i.e., have 0 phase offset)



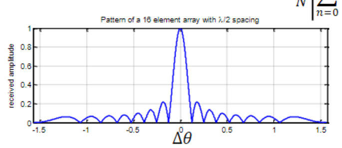
(Note: $\omega = \pi \sin\theta$)

- Antenna array gain pattern: When array is beamformed toward angle θ, what is the signal received from angle θ+Δθ?

Array response: $\mathbf{a} = [1, e^{j(\omega+\Delta\omega)}, e^{j2(\omega+\Delta\omega)}, \dots, e^{j(N-1)(\omega+\Delta\omega)}]$

Normalized weights: $\mathbf{w} = \frac{1}{N}[1, e^{-j\omega}, e^{-j2\omega}, \dots, e^{-j(N-1)\omega}]$

Antenna gain at angle $\Delta\theta = |\langle \mathbf{w}, \mathbf{a} \rangle| = \frac{1}{N}\left|\sum_{n=0}^{N-1} e^{jn\Delta\omega}\right| = \frac{\sin(N\Delta\omega/2)}{Nsin(\Delta\omega/2)}$

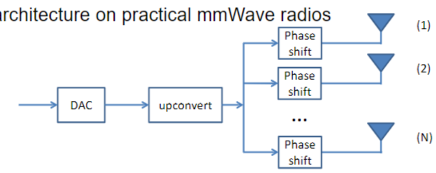


Note: $\omega = \pi \sin\theta$

$\Delta\omega = \pi\sin(\theta + \Delta\theta) - \pi\sin(\theta)$

5. How is phased-array beamforming realized in practice? How does it differ from MISO or SIMO diversity?

- Phased-array architecture on practical mmWave radios

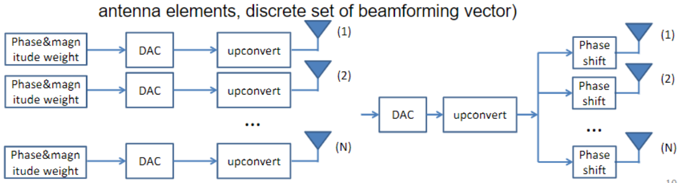


✓ Single RF chain, many antenna elements, each connected to a phase-shifter (and sometimes an adjustable attenuator)

✓ Phase-shifter and attenuator essentially assign beamforming "weights" to different antenna elements; a set of weights is called a **beamforming vector**

✓ Each beamforming vector forms a beam pattern, steering the signals towards a different "direction"

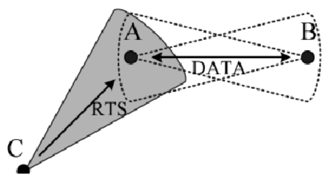
- Phased-array beamforming (analog beamforming) vs. digital beamforming
  - The MISO, or SIMO diversity gain is achieved through digital beamforming (multiple RF chains, each connected to one antenna; beamforming vector can be set arbitrarily)
  - mmWave uses analog beamforming (single RF chain, multiple antenna elements, discrete set of beamforming vector)



6. How does CSMA work in directional mmWave networks? What are the limitations?

➢ CSMA based

- Directional carrier sensing
- Open problems: hidden terminals & deafness
- Studied in ad-hoc directional MAC protocols (~2005), but more challenging due to higher directionality, more beams, and imperfect beam patterns



Deafness problem: Transmitter C keeps sending RTS to receiver A, but A cannot hear. So eventually C drops the packet after many trials.

➢ Challenge:

- A phased-array may have hundreds of beam directions to steer to
- The TX&RX must decide on the beam direction of each, to maximize "alignment", thus maximizing link SNR

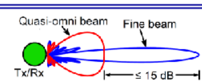
➢ 802.11ad beamforming

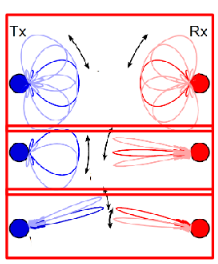
- Decision making in BHI (can be updated dynamically during DTI)
- Essentially a beam training (selection) process

7. Why is beamforming training needed in 802.11ad? Understand the high level idea of 802.11ad beamforming training

802.11ad MAC: Beamforming training

➢ Hierarchical beam training





Sector level sweeping (SLS): quasi-omni beams

Multiple sector ID detection (MID): TX quasi-omni, RX directional, or vice versa

Beam Combining (BC): both TX and RX are directional

Lecture 13. Mobile IP; routing models Why is mobile IP needed? Work flows, pros and cons of indirect routing and direct routing.
Without Mobile IP, devices must tear down and set up connections as they move from location (network) to location (network)

Mobile users don't want to know that they are moving between networks (mobility needs to be transparent to network users)

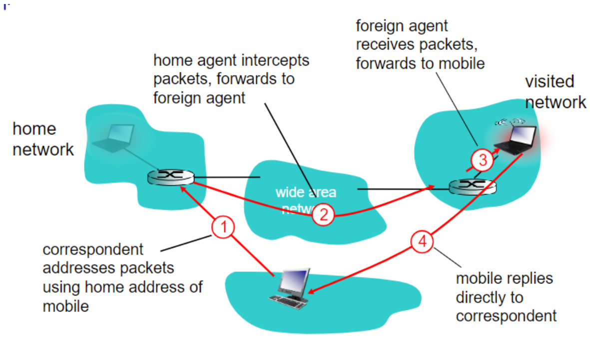
Mobile IP (Internet Protocol) is needed to facilitate seamless communication for mobile devices as they move across different networks. Traditional IP addresses are tied to specific network segments, and when a device moves to a new location, its IP address becomes outdated. Mobile IP allows a device to maintain connectivity and keep the same IP address regardless of its physical location, enabling continuous communication and mobility support.

Indirect Routing

Pros:mobility, changing foreign networks—all transparent: ongoing connections can be maintained!

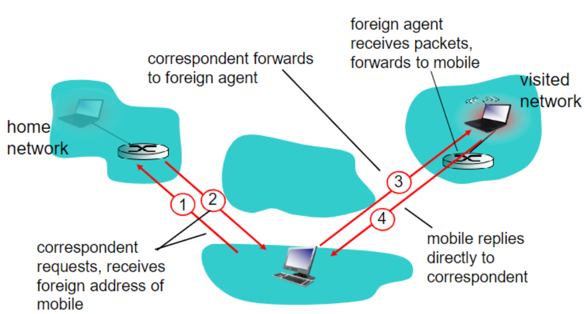
Cons:Problem: triangle routing: correspondent to home-network to mobile to... inefficient when correspondent and mobileare in same network

- Mobile uses two addresses:
  - permanent address: used by correspondent (hence mobile location is transparent to correspondent)
  - care-of-address: used by home agent to forward datagrams to mobile
- Suppose mobile user moves to another network
  - registers with new foreign agent
  - new foreign agent registers with home agent
  - home agent update care-of-address for mobile
  - packets continue to be forwarded to mobile (but with new care-of-address)



Direct Routing Pros: Overcome triangle routing problem

Cons: Non-transparent to correspondent: correspondent must get care-of-address from home agent



Lecture 14. Wireless routing

1. Formulate an LP for a given network design problem; convert LP formulation to matrix format

Standard format:

Minimize: $\sum_{j=1}^{n} c_i x_j$

Subject to: $\sum_{j=1}^{n} a_{ij} x_j \geq b_i, \quad i = 1, \dots, m$

$x_j \in \{1,2,3,\dots\}, \quad j = 1, \dots, n$

Equivalent **matrix format:**

Minimize: $\boldsymbol{c}^T \boldsymbol{x}$

Subject to: $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$

$\boldsymbol{x} \geq \boldsymbol{0}$

$\boldsymbol{x} \in \boldsymbol{Z}^n$

- Notations

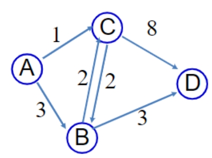
N: Set of nodes in the network

E: Set of edges

which is a subset of possible edges, i.e., $\{(i,j): i,j \in N, i \neq j\}$

b_i: Supply at node i

(demands are just negative supplies)



$x_{ij}, (i,j) \in E$: Quantity to ship along edge (i,j); a decision variable

$c_{ij}, (i,j) \in E$: Cost of shipping one unit along edge (i,j), known parameter

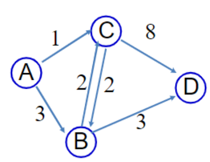
$u_{ij}, (i,j) \in E$: Capacity of edge (i,j), known parameter

12

Minimize: $\boldsymbol{c}^T \boldsymbol{x}$

Subject to: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \quad \boldsymbol{x} \geq \boldsymbol{0}, \quad \boldsymbol{x} \leq \boldsymbol{u}$

- A is called node-edge adjacent matrix; usually sparse
- Example: A, C each supplies 1 unit flow. D demands 2 units.



$x = [x_{AC}, x_{AB}, x_{BC}, x_{BD}, x_{CB}, x_{CD}]^T$

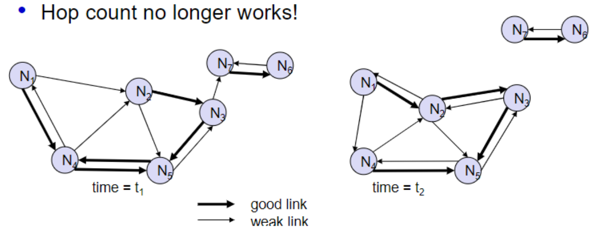
$b = [1,0,1,-2]^T \qquad c = [1,3,2,3,2,8]^T$

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & -1 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 & 0 & -1 \end{bmatrix}$$

15

2. What are the unique challenges for routing over wireless ad-hoc or mesh networks? How does DSR handle these challenges? What are the limitations of DSR?

- Challenges in wireless routing
  - Highly dynamic network topology: due to mobility
  - Highly dynamic link quality: due to channel condition changes
  - Cannot afford too frequent routing message exchange
  - Hop count no longer works!



- Reactive routing protocol to resolve the above challenges
- 2 phases, operating both on demand:
  - Route discovery
    - ✓ Used only when source S attempts to send a packet to destination D
    - ✓ Based on flooding of Route Requests (RREQ)
  - Route maintenance
    - ✓ makes S able to detect, while using a route to D, if it can no longer use its route (because a link along that route no longer works)

- The packet header size grows with the route length
- Route caching has its downside: stale caches can severely hamper the performance of the network
- Route requests tend to flood the network and generally reach all the nodes of the network
- Risk of many collisions between route requests by neighboring nodes
  ➔ need for random delays before forwarding RREQ (similarly for route reply RREP)

Location-aided routing may help reducing the number of useless control messages