Lecture 1

Layered System of Network Stack, Role of each layer.

Wireless Networks

1) Intermediate course — BTech (3^{rd} year & above)
   MTech /PhD

2) CSE /ECE eligible
   ↳ Concepts of networks with respect to protocols
   (MAC/PHY Layer of WiFi, LTE, 5G; Impact on Transport layer)

   ECE ⟶ How are these communication techniques used by computers to communicate?

3) Latest protocols and applications coming up in wireless networks
   - Localization, both outdoor and indoor
     → GPS
     → Indoor Localization ; not a mature technology
   - Activity Recognition

4) Course Prerequisites
   - Some programming → C, C++ or python
   - CSE → CN; ECE → Communication

5) Grading (Based on last time)
   - Assignments → 30 % (Programming-based)
   - Midterm → 10 %
   - 1 Quiz → 10 %
   - 1 Final examination — 10 %
   - Project → 40 % → Usually about implementing some project from papers

6) Lecture Hours → Allocated 8:30am. on Tues/ Fri
   (Will try to move it to the afternoon depending on availability)

→ Google Classroom invitation

→ all assignments & announcements will be posted

→ After each lecture, will post reading material. That should be sufficient.

→ Open book & open notes

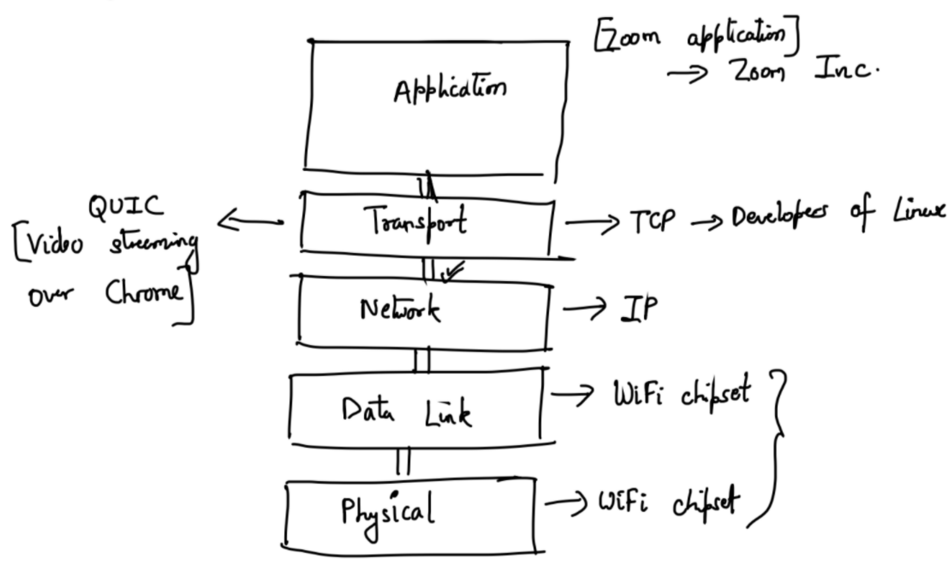
→ But no discussions among students allowed.

Wireless Networks

The network stack is divided into layers.

OSI Model → 7 layers

TCP/IP Model → 5 Layers ✓



Why do we need layers?

→ Abstraction

→ Thinking about the entire system together is difficult. Dividing them into layers helps us build protocols more easily.

→ Easier to optimize each layer separately

[Even in computer organization, we use a series of similar abstractions]

→ You might easily miss some subtle behaviors of the lower layers.

→ Abstractions are models, and have some errors.

→ Upper layers have to make some assumptions about the behavior of lower layers.

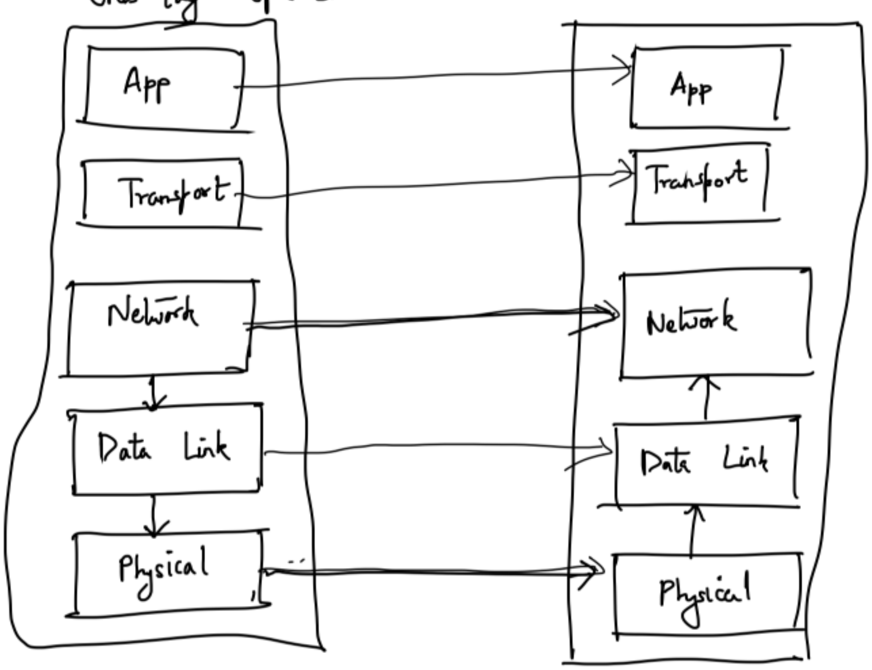
→ Division of responsibility

Data link and physical layers together can ensure that the number of packets lost is relatively small (<1%).

→ Transport layer

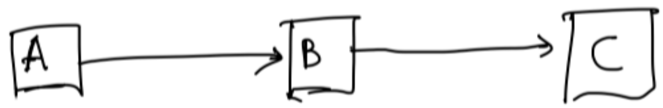
$\underbrace{\text{(TCP is the most common one)}}$ Critically depends on that assumption.

Cross-layer optimization



PHY Layer

1) Transmission of info over a single hop.



2) Radio transmission
3) Signals suffer losses as it travels through air (channel)
   → How to build a reliable communication technique in the presence of such losses?
   → For example, channel losses can be handled by retransmission

→ Send more data has some tradeoff with reliability

Data Link Layer

— Coordinating multiple transmissions over a link

— Wireless is broadcast medium; need to share the channel efficiently.

⇓

anybody who is between the transmitter and the receiver can also receive the wireless signal

— Want to avoid a case where data is sent by multiple transmitters at the same time over the same channel → no receiver will be able to decode the signal.

— Giving access to the medium to specific transmitters → different techniques of doing it.

Network Layer

— Handles routing

→ To which device a packet should be forwarded to, if it is to eventually reach its destination (IP is used)

↳ Border Gateway Protocol (BGP)

— Needs to support mobility

Transport Layer

— End-to-end transport of bytes/messages

— Needs to handle mobility

— Wireless channel is lossy; how transport layer reacts to these losses decides its performance

Wired network → (Loss of packets implies congestion.)

Wireless network → Challenging medium

Application Layer

→ Occasional disconnections

→ Sudden changes in bandwidth and latency

→ Application needs to adapt to these changes (how to do it? → Video streaming technologies)

Lecture 2:

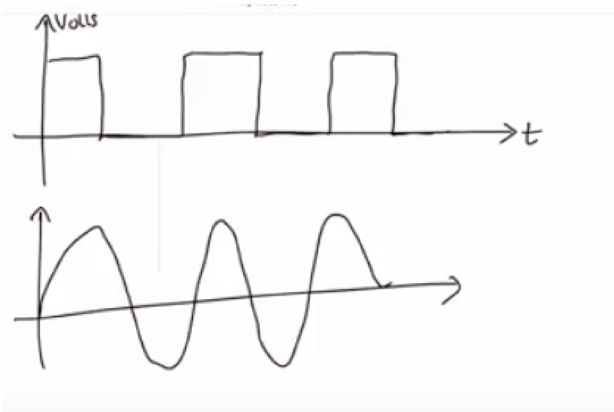
Physical Layer of wireless networks

- Log and linear units of power
- Fading -- Rician and Rayleigh models
- Shadowing
- Path loss

Reference: Sections 2.3, 2.6-2.8 of Andrea Goldsmith's book

Why are wireless networks more challenging than wired networks?
At the highest layer, applications tend to lose more packets, observe higher latency and also observe higher variations in latency, lower bandwidth and observe higher variations in bandwidth.
Sometimes depending on the situation, you can lose more intermittent connectivity.
We have to understand the root cause of these problems and quantify them.

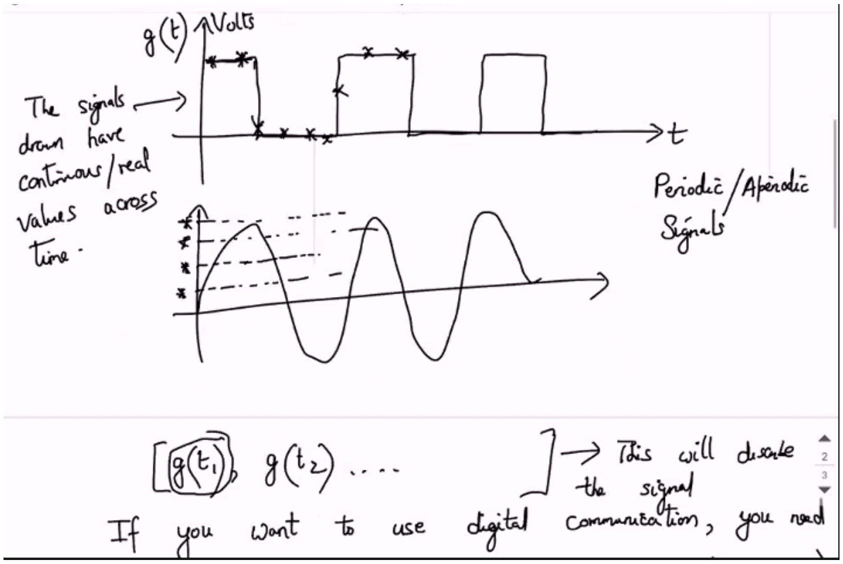
Signals:
It is a set of data that is sent. It is a function of time.



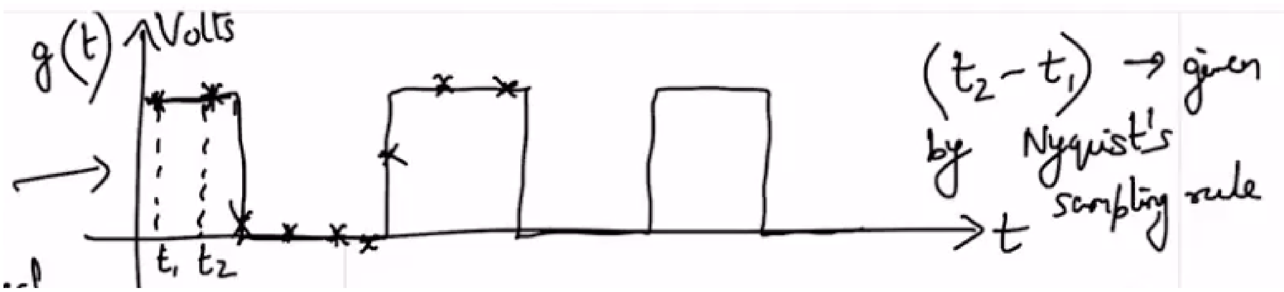
Signals may be periodic/aperiodic in nature.
We need to use digital communication in wireless networks. We sample a particular point and then send that. We denote that by g. And then we can send the vector.

Nyquist Sampling Rule

The fundamental rule in digital signals: ***If you sample the points at a rate greater than twice the frequency of signals, then there is no sampling error incurred at the receiver end*** because the signal can be exactly reproduced as a result.

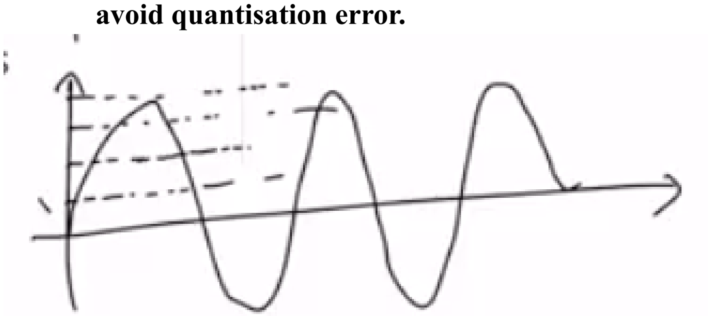


|t2-t1| = nyquist sampling rule
Effectively distance between two points is governed by the sampling rule.



If you want to use digital communication then you need to use two things.

1. Sample the points at specific points. (sampling error)
2. Quantise the data, g(t1) you take data in form of bits, you have to send one among the four values. You cannot send the actual values, bits need some amount of quantisation error. That error has to be incurred, **you cannot avoid quantisation error.**



There is another standard way of representing this in the form of complex numbers.
Assuming that the signal is aperiodic.

The **periodic signals** can be specified in a finite number of samples, periodic signals cannot carry data. They can carry only a finite amount of data simply because once you have already sent something, the amount of data you send is repeated, To send actual data, we use aperiodic signals. First, sample them and they become digitised and quantized later on.

What you get is an infinite signal in the form of

$$[\,g(t_1) \ldots\ldots \quad \ldots \quad g(t_n) \ldots\ldots\,]$$

$\downarrow$

Amplitude + Phase of the signal

This contains the amplitude + phase of the signal
Together with if you specify the amp and phase of the signal, that actually describe the entire signal completely.
Frequency can be recomputed from the amplitude.
Because we are effectively specifying the frequency at amplitude, this frequency and amplitude can be specified in two ways

$$\boxed{A\cos(\omega t + \theta)} \longrightarrow \text{Can also be represented as a complex numb}$$

The complex number form has advantages, it is easier to perform mathematical operations on them.

① Two signals

$$\left.\begin{aligned} &A_1\cos(\omega t + \theta_1) \\ &A_2\cos(\omega t + \theta_2) \end{aligned}\right\}$$

If we want to add then addition will take some trigonometric formulas but if we are using the complex form then in complex form suppose you write this as

① Two signals

$$\left.\begin{aligned} &A_1\cos(\omega t + \theta_1) \\ &A_2\cos(\omega t + \theta_2) \end{aligned}\right\} \text{ Addition would take trig. formulas}$$

$$② \left.\begin{aligned} &A_1^r + A_2^c \cdot j \\ &A_2^r + A_2^c \cdot j \end{aligned}\right\} \text{ Addition just requires adding two numbers.}$$

Usually, signals are represented as complex numbers
This real form in this case is

$$A_1^r = A_1\cos(\omega t + \theta_1)$$
$$A_2^c = A_1\sin(\omega t + \theta_1)$$
$$A_1 = \sqrt{(A_1^r)^2 + (A_1^c)^2}$$

A cellular network is also a wireless medium, Bluetooth is also a wireless medium. All of them use this concept.

We will get into the ***differences between wired and wireless networks.***
At the physical level, both are similar, they are using some sort of modulation coding scheme to send the data. **The key difference** is **how they are managing this access to the medium**. You need to make the decision that this is the broadcast medium, if everyone sends data at same time, how will you avoid collisions?

The **main difference** is how they are managing this access to the **air medium**.

In *cellular networks* the **access is centralised.** There is a master-slave architecture. But in **WIFI, it is decentralised,** there is no central entity that is controlling. The primary difference lies in the data link layer, also wifi uses less power and its transmission range is small.

Signal energy and power

Signal energy is E_g

Signal energy and power : →

$$E_g = \int_{-\infty}^{\infty} |g(t)|^2$$

E_g divided by Time is power.

Power is measured in mW or W when we want to measure in linear units. It is also often measured in log units. We use dBm. The concept of dBm is important here . ***dBm is decibels milliwatts.***

use dBm (

$$P_{mW} = 1\,mW \times 10^{P(dBm)/10}$$

After taking the logarithm, we will get

$$P(dBm) = 1.2589\ P(mW)$$

P(dBm) = 1.2589 P(mW)

dBm means decibel mWs. We are more familiar with it like **1000 mW = 1W.**

Power in the wireless transmission is very less.
Even in cellular, cell towers don't use more than **30dBm or 1W.**
Wireless Networks are challenging and there is a reason for that.
What is challenging is we measure the **ratio of power transmitted to power received**?
If you want to measure this, denote it by log(Pt/Pr) you mean log(pt)-log(pr)
Path loss is the measure of the amount of channel distortion.

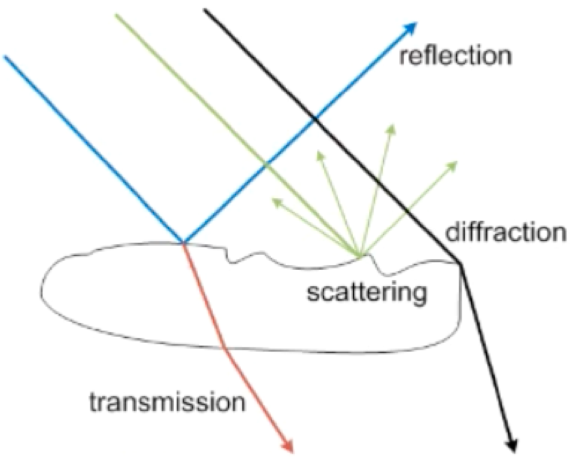
There are different types of distortion

Path Loss is the ratio of power transmitted to power received. More like the good put. The channel distortion is measured in logarithmic scale instead of ordinary scale that too in dBs.

1) What is the ratio of power transmitted to power received? $\longrightarrow$ Path loss

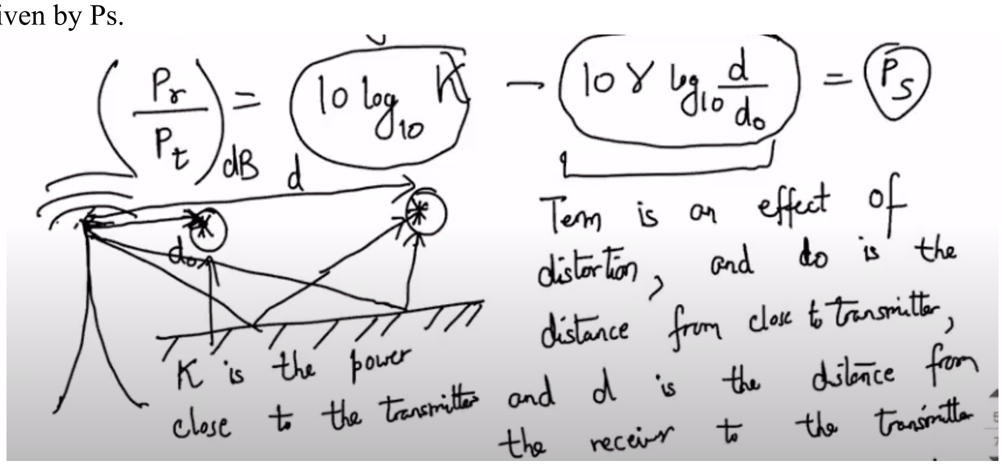
$$\log \frac{P_t}{P_r} = \left[\log P_t - \log P_r\right] \text{dB}$$

Measure the amount of channel distortion



In wireless networks, the signals encounter obstructions due to the presence of ***buildings,vegetations*** etc.

1. **We need to compute the effect of the obstructions.**
    a. Obstructions lead to ***partial reflection, refraction diffraction and absorption,scattering also***.
    b. It is very hard to know, much of what is happening. If you want to utilise, **you do ray tracing to identify the impact on signal propagation.**
    c. This can be very complicated in practice because it is very obvious that it is very computer-intensive. It is still be done using GPU's in controlled settings , but it is very rare.

2. The **second approach** is instead of *directly quantifying*, you can use a **statistical model to quantify the impact.** You can use a **lognormal random variable** to model this with **mean 0** then the power received becomes (in dBm) given by Ps.
    a. 

    b. d_0 is the distance from close to the transmitter

c. And d is the distance from the receiver to the transmitter.
d. γ is the path loss exponent depending on the specific propagation environment.
e. κ *is the power close to the transmitter*
f. $-10\gamma\log(\frac{d}{d_0})$ is the term which is an effect of distortion.

$$\begin{aligned} P_r[\mathrm{dBm}] = P_t[\mathrm{dBm}] & \\ & \boxed{+\,\mathrm{pathloss}[\mathrm{dB}]} \\ & +\,\mathrm{shadowing}[\mathrm{dB}] \\ & +\,\mathrm{multipath}[\mathrm{dB}] \end{aligned}$$

the corresponding phase shift relative to the LOS path is approximately

$$\Delta\phi = \frac{2\pi\Delta d}{\lambda} \approx \frac{\pi}{2}v^2,$$

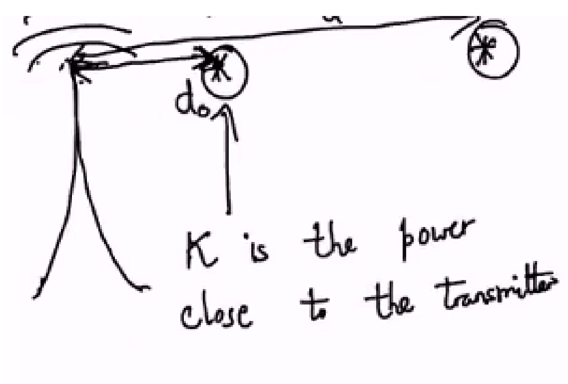
The dB attenuation is thus

$$P_r \text{ dBm} = P_t \text{ dBm} + K \text{ dB} - 10\gamma\log_{10}\left[\frac{d}{d_0}\right].$$

K is the unitless constant that depends on the antena characteristics and the average channel attenuation.

$K\ dB = 20log_{10}\frac{\lambda}{4\pi d_0}$

Path loss can be treated separately from shadowing.
Shadowing is the large scale random variation around the path loss. Log normal model is quite good.
The statistical model implies that you are no longer trying to quantify the impact from the fundamentals. You are trying to observe and then quantify based on observation only.
What can you observe to quantify this?
Effectively you have to measure, you cannot very directly measure exactly how much power is transmitted, you think it of like an antenna system, or consider it as a Wifi tower. What will be the power transmitted, we don't know.
The power that you are transmitting is more than the power received by the receiver.



This type of distortion due to obstructions is called shadowing. (LONG TERM FADING)

In cities, it is common due to the presence of very tall buildings.

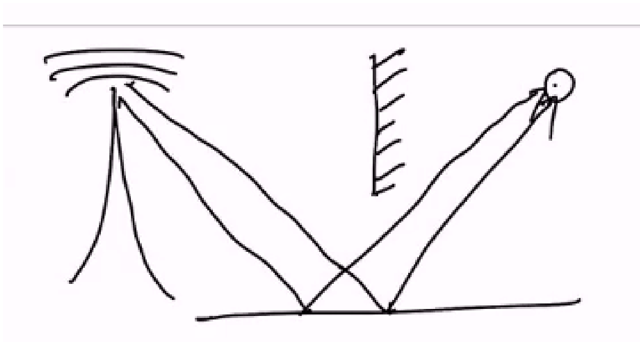
Short term/ Small scale FADING
Multipath means, if you have this transmitter and receiver and suppose you have some height obstructions, the signal will go with different paths. When we are adding, the phase, we will get like a lot of signals.

$$A_1\cos(\omega t + \theta_1) + A_2\cos(\omega t + \theta_2) + \cdots$$

Fading can spread out the signal thus introducing some short term distortion. The distortion is not linear, it is low somewhere, high somewhere, distortion is changing over time. Usually to denote this fading, Ps - shi(fading)

There are formulas to model shi

Suppose if you have a line of sight, then in major cities, you have a specific model called the **Rician model of fading(LOS)**, and if you don't have any line of sight, the **Rayleigh Model(w/o LOS)**.



Here we don't have any line of sight, a different model of fading known as the **Rayleigh model.**
The key difference between Rician and rayleigh) is Rician model has one dominant component of the signal other than the components obtained via reflection. Apart from this at the receiver end, there is some noise.

3. **Receiver Noise:**
   a. Some noise at the antenna and sensing system is usually modelled as a zero-mean Gaussian random variable with fixed variance.
   b. Suppose all these things were not involved, there are no obstructions, no fading, the receiver is perfect, even then you will lose some power, because of path loss, due to distance.
   c. We tend to have some path loss due to distance.
   d. Even in space or atmosphere some amount of the power is lost due to distance. This can be measured using the path loss coefficient.
   e. There is a path loss coefficient to quantify based on distance how much power is lost. So all this together makes up the amount of distortion faced by the channel.
   f. Path loss will come irrespective of the shadowing. Path loss is relatively is a short factor than fading and shadowing, higher frequency path losses come to show significantly.

Lecture 3:

- Need for modulation
- Utilization of different frequency bands
- Different modulation techniques
- Fourier Transform

Concepts of fading along with modulation make numerical complete.
We were talking about the properties of transmission over a wireless network. Information can only be sent through aperiodic signals. Only aperiodic signals have that element of surprise. That element of surprise or unexpected signal can be used to encode the data.

Transmission of Practical Signals:

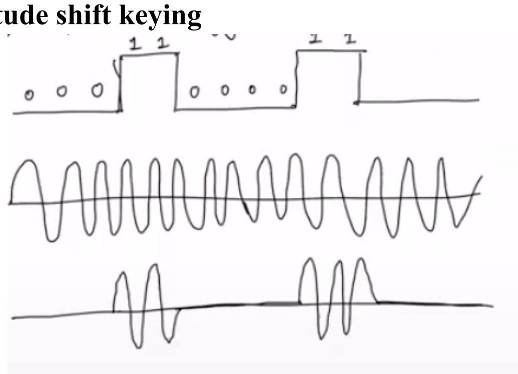
1. The signal encoding the data(baseband signal) is never sent directly. When you are transmitting, you need a receiver at the receiver end.
2. The antenna length should be of the order of the wavelength. If our data has a relatively low frequency. Frequency = c/wavelength.
    a. $\lambda = \frac{c}{f}$ → Low in baseband signals
    Then, we need λ to be very high, which means we need very
    b. If the frequency is very low, then we need wavelength to be very high, which means we need a very high antenna.
    c. This can only be avoided by mixing the original data signals with the periodic high-frequency signals. This process of mixing is called Modulation.

3. By using modulation, it is possible to send multiple streams of data in parallel because you are mixing it with a high periodic signal, it is called the **carrier signal**. By different carrier signals, we can use multiple streams of data.
    a. Suppose 4G and Wifi are being used together, without any interference, as there are working at different frequencies. They use different carrier waves.
    b. When we tune, TV/Radio channels are switching the frequencies of receiving signals.
4. Allows more information to be encoded. More than one baseband signal can also be potentially added. In this way, modulation is being used everywhere. No communication system avoids modulation. This carries some interesting implications.
5. **Every wireless communication technology needs to use some type of carrier wave**. Even a single device can use multiple carrier waves. There are also some ways to send using one carrier wave also. When you are talking, you are being shared with other people.
6. Wifi Access Point can encode data for multiple phones on the same carrier, each of the clients will understand in their own way, this technique has some implications, each wireless technology is diverse these days.
7. We have, Wifi, LTE, GPS, Bluetooth, NFC etc. They use different types of carrier waves. All of these techniques can work in parallel. If multiple signals are sent with the same frequency at the same time within the range of each other, then we will see, collisions happening. And in case of collision, none of the recipients will be able to read the data. That is why., what has been done,
8. There is a govt. Regulatory authority decides which technology can use which frequency band
    a. TV works in UHF, VHF band, less than 100 Mhz
    b. 700-900 Mhz for smartphones
    c. 1.532 GHz for GPS
    d. 2.4 GHz for Wifi/Bluetooth/ZigBee, 5-6 GHz Wifi.
    e. Beyond that LTE
    f. 22-26 GHz→ 5G communication

9. Initially, wifi started with the 2.4 GHz band, but then it was very popular, the usage was not supported by the limited band. At 5GHz, it was not used that much, but for vehicular communication, we used a lot in 5-6 GHz. But it turned out to be given to Wifi and now Wifi used 5GHz band
10. **Capability to penetrate walls decreases when frequency increases**. The recent technologies that use higher frequencies require more dense base stations.
11. As Wifi/Bluetooth works at the same frequency band, then they use some techniques at the data link layer to avoid a collision. CSMA/CD
12. Who is enforcing these regulations? In India, we have the telecom regulatory authority of India TRAI
13. → Users in the vicinity would observe loss of service
    → That complaint by the service providers would be forwarded to TRAI.
14. What if somebody is not following regulations?
    → Possible, but some sophisticated hardware is necessary.
15. There have been some spoofing attacks that can be launched using these desires.
16. How is this technology enforced, by some licence. Based on this licence through software defined radio, they are allowed, we are allowed to transmit within some range, that frequency is not licenced to anyone. It is possible to spoof GPS signals using it. **Your phone is hacked, your wifi cannot transmit 4G, as it is hardware bound, it cannot be tampered by any software intrusion. If you want to do experiment, you would like to define software-defined radio, in actual devices it is encoded in the hardware itself.**

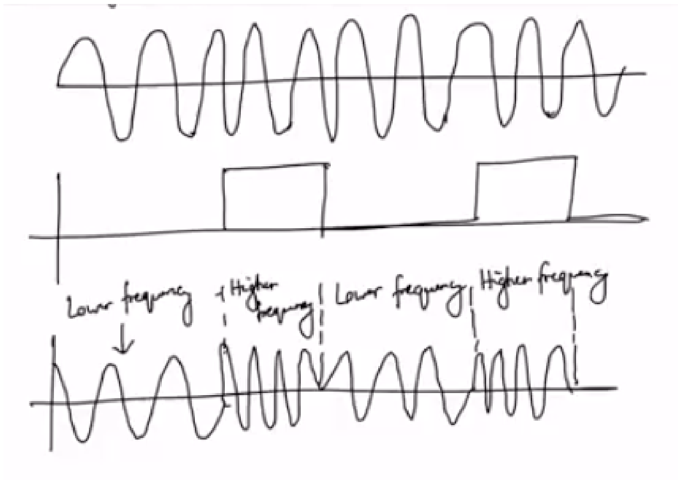
- pulse amplitude modulation (MPAM) – information encoded in amplitude only;
- phase-shift keying (MPSK) – information encoded in phase only;
- quadrature amplitude modulation (MQAM) – information encoded in both amplitude and phase.

Mixing of signals:
There are a number of ways you can mix signals, like

1. **Amplitude shift keying**
   a. 

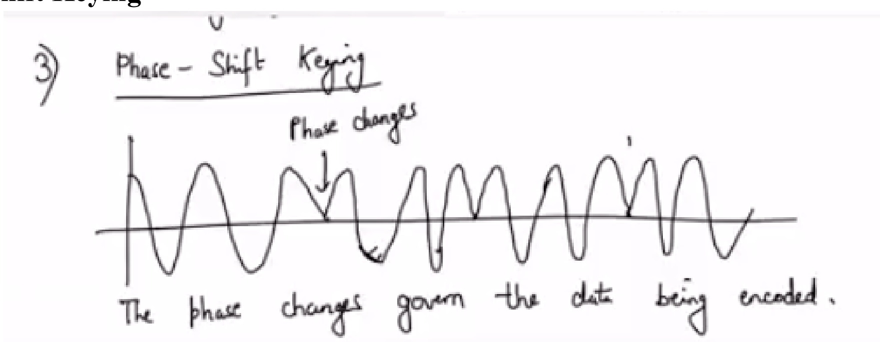
   b. If you denote the baseband signal as m(t) and carrier signal as $A\cos(\omega_c t)$ and modulated signal is represented as $Am(t)\cos(\omega_c t)$

2. **Frequency shift Keying**

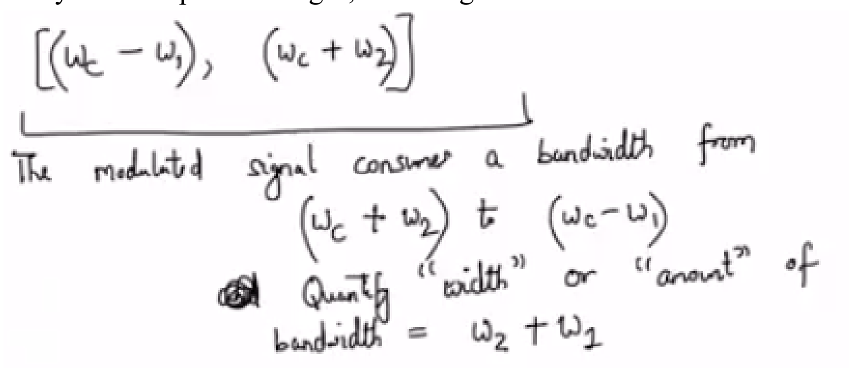
a. 

b. The frequency governs the value of the signal.

3. Phase Shift Keying

a. 

b. Every time the phase changes, we change 0 to 1 and 1 to 0.

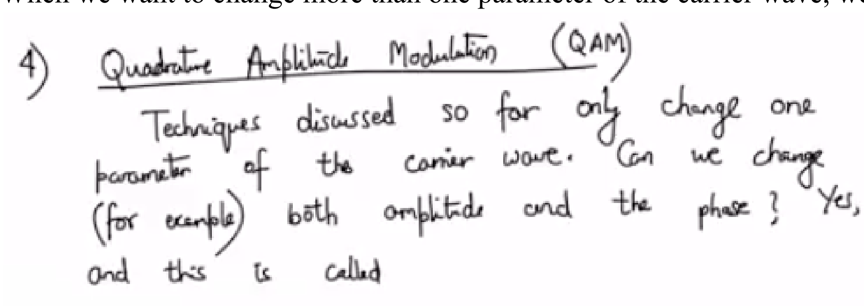
c. 

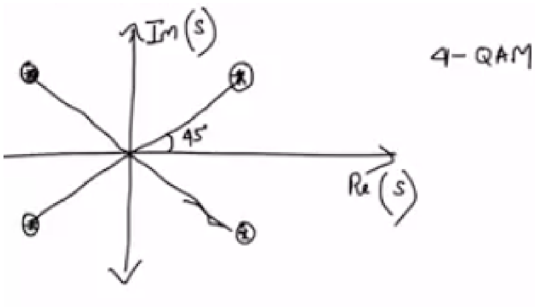
d. Anybody transmitting at the same time in this frequency range will see the collision

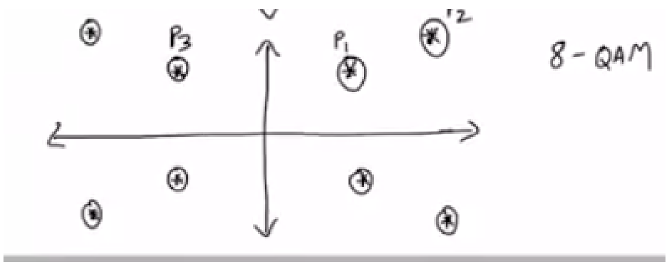
e. The carrier wave is the first pick and after mixing, we got the modulated signal.

4. Quadrature Amplitude Modulation

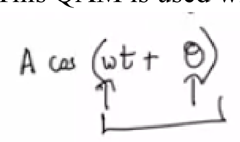
a. When we want to change more than one parameter of the carrier wave, we use QAM.

b. 

c. 

d. 

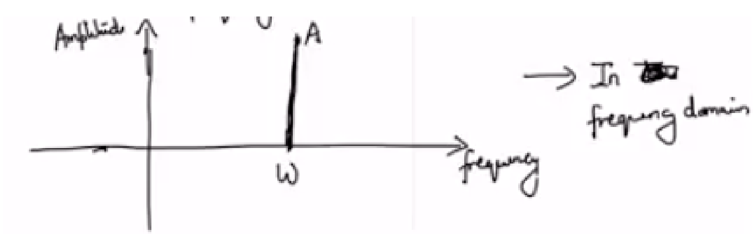
e. P1 and p3 have the same amplitude but different phases, P1 and P2 have the same phase but different amplitude.
f. This QAM is used widely in practical systems.

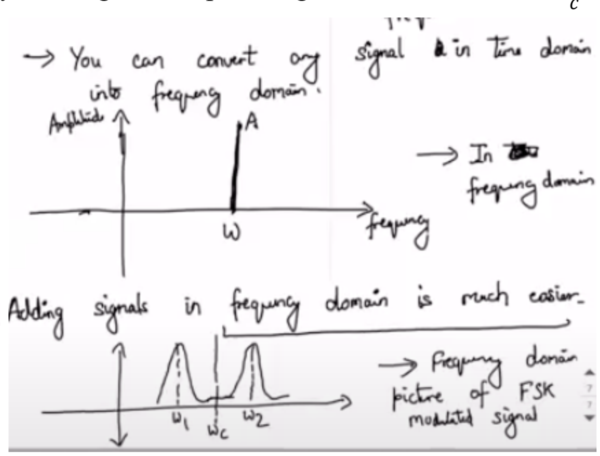
g. 

Fourier transform:

If you have any periodic signal, you can represent it as the sum of sinusoids. This thing has lot of implications, whatever the signal you have, it will always have mixture of different frequencies and it used the technique of fourier transform and the idea is you can actually convert any signal from time domain to frequency domain.

When I have Acos($\omega t + \theta$), effectively this is the time domain because t is the parameter there. In the long run, not only t is the parameter, if I wanted to plot it in terms of frequency, I will get an impulse signal here. It is called ω_c





All the signals are infinite data series, the advantage to convert into frequency domain is, addition of signals is very easy in frequency domain.

Adding the signals in time domain is very difficult, so first, we convert any given signal in time domain into frequency domain. Now the addition in the frequency domain becomes relatively easier.

Lecture 4

- Impact of Fading and its Quantification
- Doppler Shift and Impact of Mobility
- Types of multiplexing

3.2.2 Envelope and Power Distributions

For any two Gaussian random variables X and Y, both with mean zero and equal variance σ^2, it can be shown that $Z = \sqrt{X^2 + Y^2}$ is Rayleigh distributed and that Z^2 is exponentially distributed. We have seen that, for $\phi_n(t)$ uniformly distributed, r_I and r_Q are both zero-mean Gaussian random variables. If we assume a variance of σ^2 for both in-phase and quadrature components, then the signal envelope

$$z(t) = |r(t)| = \sqrt{r_I^2(t) + r_Q^2(t)} \tag{3.31}$$

is Rayleigh distributed with distribution

$$p_Z(z) = \frac{2z}{\bar{P}_r} \exp\left[-\frac{z^2}{\bar{P}_r}\right] = \frac{z}{\sigma^2} \exp\left[-\frac{z^2}{2\sigma^2}\right], \quad z \geq 0, \tag{3.32}$$

where $\bar{P}_r = \sum_n \mathbf{E}[\alpha_n^2] = 2\sigma^2$ is the average received signal power of the signal – that is, the received power based on path loss and shadowing alone.

We obtain the power distribution by making the change of variables $z^2(t) = |r(t)|^2$ in (3.32) to obtain

$$p_{Z^2}(x) = \frac{1}{\bar{P}_r} e^{-x/\bar{P}_r} = \frac{1}{2\sigma^2} e^{-x/2\sigma^2}, \quad x \geq 0. \tag{3.33}$$

Thus, the received signal power is exponentially distributed with mean $2\sigma^2$. The equivalent lowpass signal for $r(t)$ is given by $r_{LP}(t) = r_I(t) + jr_Q(t)$, which has phase $\theta = \arctan(r_Q(t)/r_I(t))$. For $r_I(t)$ and $r_Q(t)$ uncorrelated Gaussian random variables we can show that θ is uniformly distributed and independent of $|r_{LP}|$. So $r(t)$ has a Rayleigh-distributed amplitude and uniform phase, and the two are mutually independent.

EXAMPLE 3.2: Consider a channel with Rayleigh fading and average received power $\bar{P}_r = 20$ dBm. Find the probability that the received power is below 10 dBm.

Solution: We have $\bar{P}_r = 20 \text{ dBm} = 100$ mW. We want to find the probability that $Z^2 < 10 \text{ dBm} = 10$ mW. Thus

$$p(Z^2 < 10) = \int_0^{10} \frac{1}{100} e^{-x/100}\, dx = .095.$$

When can we say that we have a rayleigh distribution?

If we have two random variables X and Y, with both having 0 mean and equal variance σ^2, then it can be shown that $Z = \sqrt{X^2 + Y^2}$ *is rayleigh* and Z^2 is exponentially distributed.

Power received based on path loss and shadowing alone:

where $\bar{P}_r = \sum_n \mathbf{E}[\alpha_n^2] = 2\sigma^2$ is the average received signal power of the signal – that is, the received power based on path loss and shadowing alone.

Thus, K is the ratio of the power in the LOS component to the power in the other (non-LOS) multipath components. For $K = 0$ we have Rayleigh fading and for $K = \infty$ we have no fading (i.e., a channel with no multipath and only a LOS component). The fading parameter K is therefore a measure of the severity of the fading: a small K implies severe fading, a large K implies relatively mild fading. Making the substitutions $s^2 = KP_r/(K+1)$ and $2\sigma^2 = P_r/(K+1)$, we can write the Rician distribution in terms of K and P_r as

So we got to know that $2\sigma^2$ *is exponentially distributed* and K is the power is Line of sight component to the power in the other multipath components. If we have K=0, in that case, we get rayleigh fading(Non LOS) and if $K = \infty$ then we have no fading, and LOS is clear and no multipath component. Small K means a lot of fading and a relatively bigger K means less fading. Practically this K is never 0 and infinite, it is somewhat in middle.

*Recall of **fading** concepts*

1. Log and linear units → Fading and path loss is considered in log units

$$P_r(dBm) = P_t(dBm) + K - 10\,\gamma \log\left(\frac{d}{d_0}\right)$$

P_r: At a particular Location
P_t: Power Transmitted
K: Path loss due to distance + antenna's properties
γ: Shadowing coefficient
d_0: reference distance close to the antenna

   a.
   b. Formula to calculate received power.
2. Whatever power, we have has to be convered to dBm units. Suppose we are given that $P_t = 1mW$, the parameter k=-31.59dB, $\gamma = 3.71$, $d_0 = 1m$. Find the received power at a distance of 150m? How can you find it?
   a. First thing is that you need to convert 1mW to dBm
   b.

$$P_t(dBm) = 10\log_{10}\frac{P_t(mW)}{1\ mW}$$

$$= 10\log_{10}\frac{1}{1} = 0\ dBm$$

   c.

$$= 10 \log_{10} \frac{1}{1} = 0 \text{ dBm}$$

(2) $$P_r(\text{dBm}) = P_t(\text{dBm}) + K - 10\gamma \log_{10} \frac{d}{d_0}$$

$$= 0 - 31.54 - 10 \times 3.71 \times \log 150$$

$$= 0 - 31.54 - 37.1 \times [2\log 5 + \log 3 + \log 2]$$

$$- 31.54 - 37.1 \times [2 \times 0.66 + 0.47 + 0.3]$$

$$= \ldots \text{ dBm}$$

d. Power received value
e. Gamma part is the attenuation.
f. Power received value
g. Power received value

Buy a GPS receiver $\rightarrow$ active antenna with a gain of 28 dB.

$$P_r = -100 \text{ dBm} + 28 = -72 \text{ dBm}$$

3. Wireless signal propagation is different than wired transmission.

Performance of digital modulation technique criterias
1. Probbility of error : relative to symbol or bit error
2. Outage Probability : probability that the instantaneous SNR falls below a certain threshold.

Flat fading increases either of the above two criterias
Wireless channels can cause fading as well as doppler shift.

We will discuss about the impact on digital modulation performance of noise, flat fading, frequency selective fading and doppler shift.

SNR

$$\text{SNR} = \frac{P_r}{N_0 B}.$$

6.1.1 Signal-to-Noise Power Ratio and Bit/Symbol Energy

In an AWGN channel the modulated signal $s(t) = \text{Re}\{u(t)e^{j2\pi f_c t}\}$ has noise $n(t)$ added to it prior to reception. The noise $n(t)$ is a white Gaussian random process with mean zero and power spectral density (PSD) $N_0/2$. The received signal is thus $r(t) = s(t) + n(t)$.

We define the received SNR as the ratio of the received signal power P_r to the power of the noise within the bandwidth of the transmitted signal $s(t)$. The received power P_r is determined by the transmitted power and the path loss, shadowing, and multipath fading, as described in Chapters 2 and 3. The noise power is determined by the bandwidth of the transmitted signal and the spectral properties of $n(t)$. Specifically, if the bandwidth of the complex envelope $u(t)$ of $s(t)$ is B then the bandwidth of the transmitted signal $s(t)$ is $2B$.

Mobility: Why does mobility makes a difference?

There is a physical phenomenon called doppler shift.

Doppler shift:

Whenever either the transmitter or receiver is moving, then the signal's frequency appears to change. The way it changes is there is a specific formula

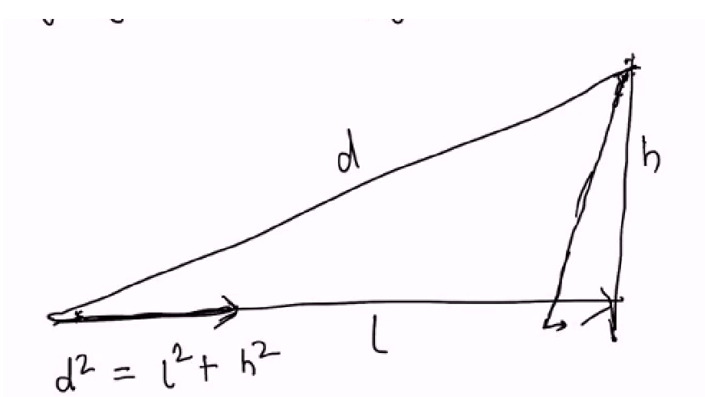
Doppler shift $= \dfrac{\Delta v}{c} * f$; c is the speed of light ; f is the frequency of the signal

Doppler shift is usually significant in practice. If you are moving in a vehicle. There is another aspect about the doppler shift that $\Delta v \;=\; \textit{relative movement between the transmitter and Receiver}$

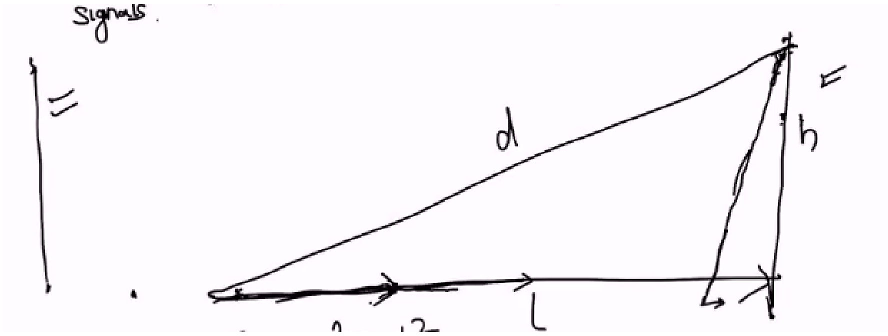
It depends on how they are moving with respect to each other. If both of them are moving in the same speed, then the relative velocity is 0, there is no change in the doppler shift.

If they are coming towards each other, then the frequency tends to increase. Then doppler shift increases the frequency of the received signal. Otherwise it reduces the frequency.

Here is an interesting aspect, in practice what happens is when we are travelling by car and trying to receive the cellular network signals. Antenna is at a height. If you are at a relatively higher distance, then at this point, your distance when you are actually close, then h will dominate. So the amount of change in d will be relatively small. But when you are far way then l dominates and so the doppler shift is higher.



All these aspects play a role that how doppler shifts actually deals with cellular networks. In cellular networks, doppler shift plays an important role. If you are standing signals are good,
If there are two cell towers instead of 1, then at some point, this signal will become weaker, at some point, this cell tower will decide, it is better to use the right one, but not left one.

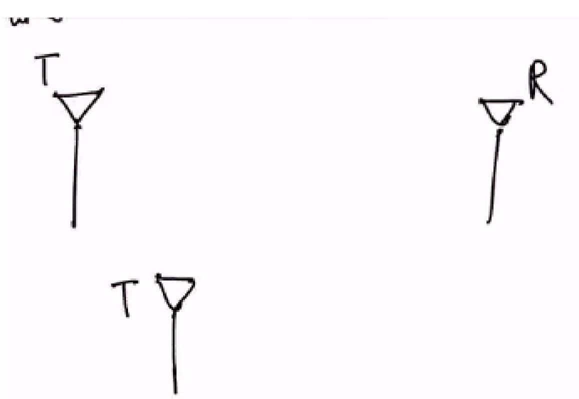


Each tower sends a special type of pilot signal, these pilot signal is understood by phone that which frequency the power is coming. In cellular networks, there is something called as pilot singal, it understands the amount of power received and the frequency at it is received. This pilot signal is crucial, it gives understanding to the smart phone, that what is the real situation w.r.t different cell towers. The distance between these cell towers t1 is smaller than t2.
When we do handoff, it leads to a very short duration of disconnection. That is one of the major problem that there is no way to know user will come back or keep going.

Assume there is a user who is just patrolling around, and the user will feel that the quality of signal is very poor.
For each technology, you have separate receiver antena and it is a hardware thing. Your phone only reports to the cell tower, we get to know.
You have Wifi/4G(Volte)
In 4G, it is a centralised system(handoffs are decided not by the phone, but by the cell towers)
It is an important concept. Wifi is exactly opposite. It is a decentralised system. Here your phone decided if you want to handoff.

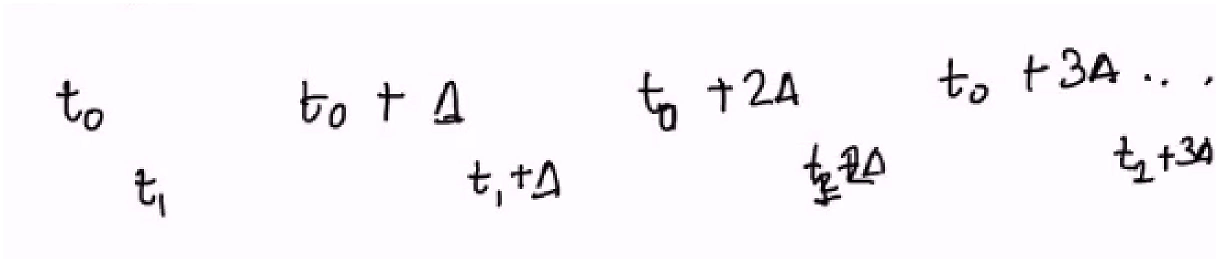
Multiplexing: It is closely related to modulation.
In general we have a lot of data over the wireless network to send. But RA spectrum is limited, so we want to pack more and more data together over the limited spectrum. As far as possible we are trying to use parallel transmission i.e. where multiple devices can send and receive at the same time.
Suppose you have a transmitter, and a receiver, and you have another transmitter and receiver.

Can they talk in parallel?

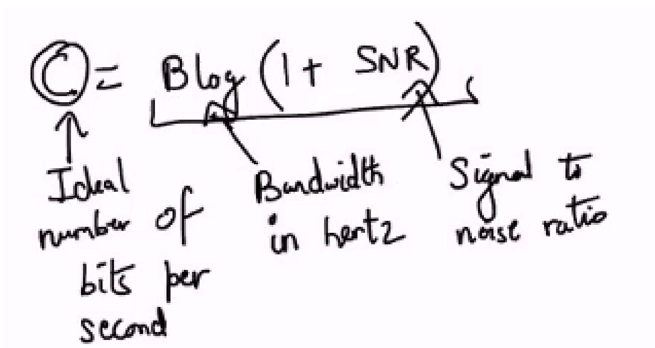


We can use Time Division Multiplexing, whenever we are sending the digital data, then we can divide up the time between the two transmitters, so that both can send together. This is very widely used again in cellular networks. Because cellular networks need to support a lot of users.

The service provider usually acquires some licence for the spectrum and then uses it via cell towers.
As the spectrum is very expensive, and a lot of people use this same spectrum. It has been observed that the human voice is not distorted, if it is sampled at some frequency (which is indeed given by Nyquist formula.). So it can divide and send multiple human voice signals at the same time.



We can multiplex as many users as given by Shannon capacity formula.
This can be very well utilised in practice.
Shannon capacity formula is Blog(1+SNR) = C = ideal number of bits per second.



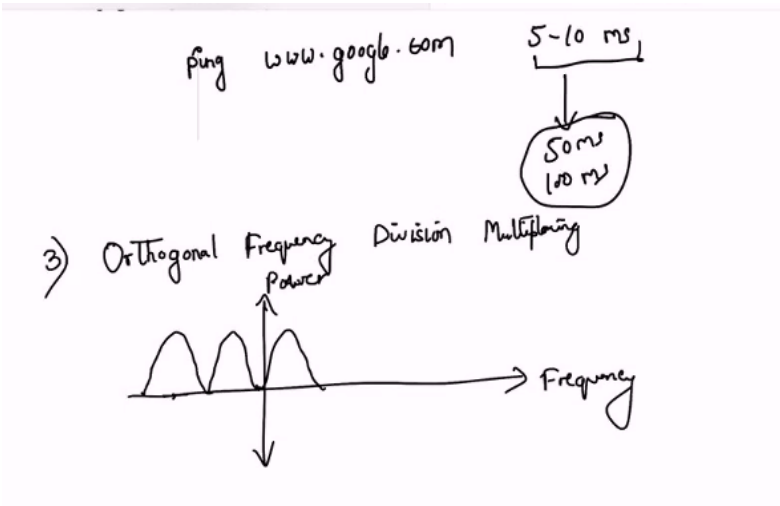
Fundamental principle is if you allocate this Bandwidth, whatever optimisation you do , you cannot carry more than C number of bits per second.
Shannon's formula is a fundamental limit on the amount of data you can send per second.
Whenever you want to understand the capacity of the Wifi signal/Cellular network, we will come to the shannon's capacity formula, we will use it again and again.
The SNR is within some log factor. If you increase the P_r then SNR will also increase. Increasing SNR at some point tends to give diminishing returns. You might need other optimisations.

Frequency Division Multiplexing:

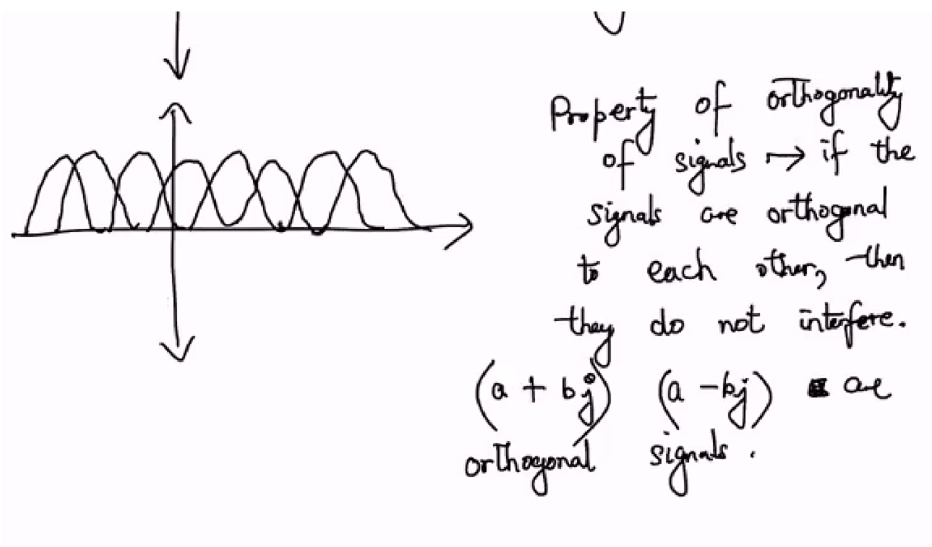
This is relatively simple. Usually, the signal is split into multiple bands, even after getting the licence, usually not all the devices use the entire bandwidth. The bandwidth is split up and allocated to different users. You can see this in Wifi, you have channel1 to channel10(2.4GHz Wifi). Within this 2.4GHz overall band, a device has to select any one of the channels with some bandwidth of 20MHz. Other devices can use some other band within this. This 2.4GHz has around 10 channels, and they can all transmit and receive using these 10 different channels parallelly. There is no requirement of any coordination once the channels are decided.
If you look at the time-division multiplexing, time coordination was necessary. Clocks are inaccurate.

FDMA is used in Wifi
In cellular networks, the tower decides.
10 channels mean only 10 users can simultaneously browse the internet. It is not that you are always fetching the data. There is ample scope that time is left. In that time
Because there is no coordination, you will lose the packet.
It will increase the latency, there will be collisions. If the space you are at is crowded, then the speed of the networks becomes slower.
Ping latencies are in the range of 5 to 10 ms from IIIT Delhi.
It is all happening at the same time.
Any data can be sent using any frequency in any of the channels.

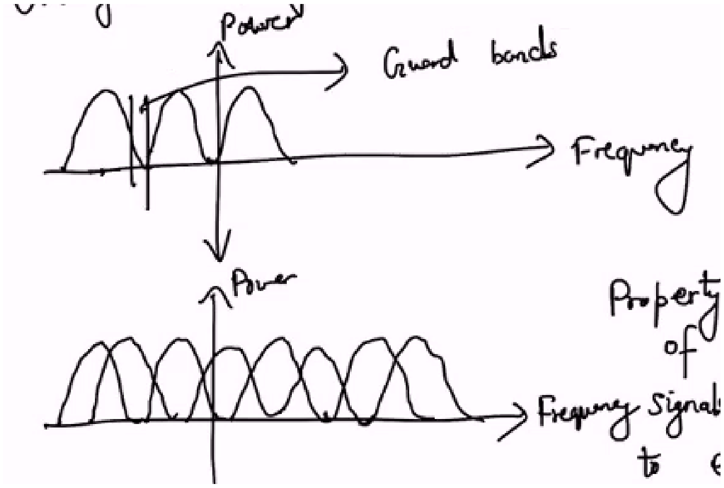
Orthogonal Frequency division multiplexing(OFDM)



There is no interference between signals, so you do not need this much gap. There is a property of orthogonality of the signals. This means if the signals are orthogonal to each other then they do not interfere.
We can express signals in the form of complex numbers (a+bj) and its conjugate is (a-bj) will be the orthogonal signals.
If you use orthogonal carrier wave you can do something like this

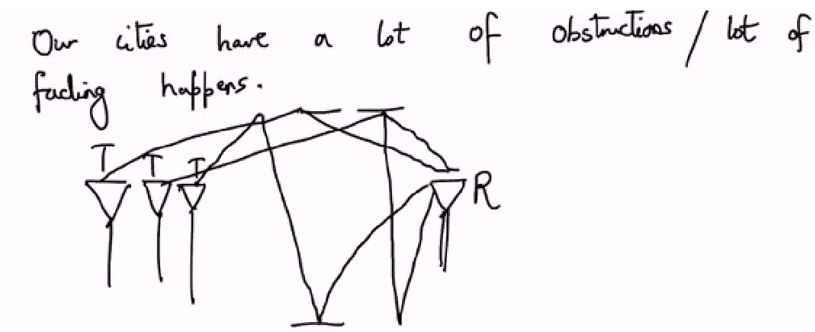


They can be used without interference.
In polar coordinates they are perpendicular.
How the overlap is happening?
It comes from the property of complex numbers. These are not in the real planes, these properties of complex numbers allow them to work together.
Can we save the Bandwidth in terms of FDM?
Yes we can but still we have disadvantages. If we have a doppler shift, then there will be more and more.

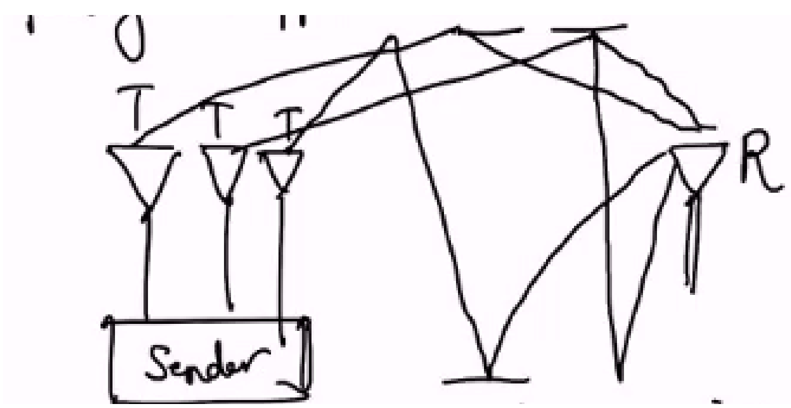


Till 2G, There was also one Code division Multiplexing and this was the technique. Wifi has these standards 802.11b, 802.11g, 802.11n

Our cities have lots of obstructions. Lots of fading happens and we mentioned this.
There is almost never a line of sight between the cell tower and the device. Isn't it possible that if we have T and R, instead of one antenna, if I use multiple antennas? The same device uses multiple antennas. These antennas are at some distance. The channel will go some reflection and so on.

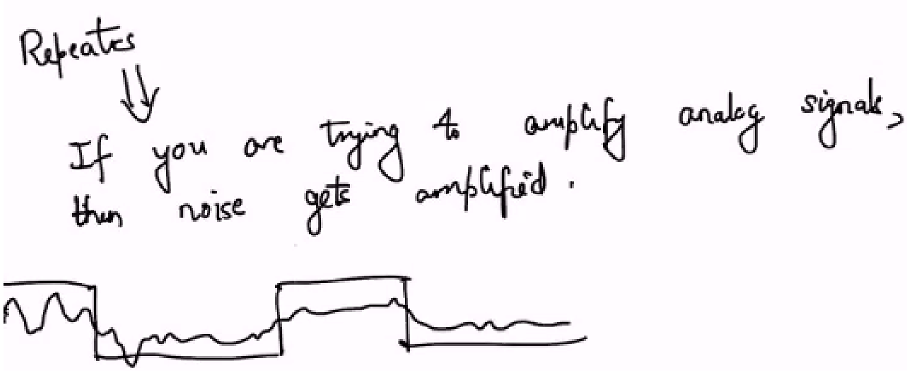


This is **Transmit Diversity.**
The receiver will receive the signals from each of these antennas after traversing different paths. Statistically, it is possible at the receiver end to separate out each of the signals even if they are at the same time and same frequency due to different phases and amplitude. Because they are coming through different paths, the receiver gets that antenna 1 comes with different amplitude and phase, even though they are at the same frequency, the receiver will look into these path characteristics. The sender is the same guy

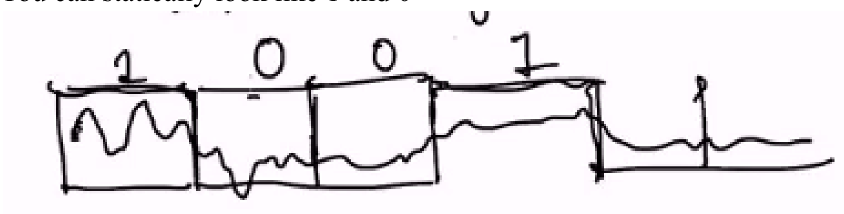


The receiver would receive the signals from each of these antennas after traversing different paths. Statistically, it is possible at the receiver end to separate out each of these signals(due to difference phase and amplitude) even if they are at the say time and frequency. These can be proven statistically also. It is widely used in 4G and beyond systems. 5G has not come in India yet. This is called Spatial Multiplexing. When you go to buy Wifi AP/Routers. It says it supports MIMO(Multiple Input Multiple Output) This is the intuitive idea. Fundamentally you cannot do, other than have more dense transmitters, on the other side at the receiver end you can have amplifiers. There are some devices in the digital networks known as repeaters. If you try to amplify the analog signals, then the noise also keeps increasing.

Because digital communication uses quantisation, then you will be able to reconstruct the signals using repeaters.



You can statically look like 1 and 0



If your noise is not that high, you can reconstruct it. You can handle it by a more dense network here.

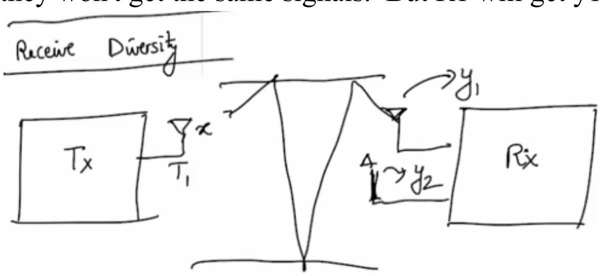
Lecture 5

- Fundamentals of (point-to-point) MIMO

- Transmit and Receive beamforming
- Tradeoffs involved between SNR and utilization of space-time coding

We discussed time, frequency, orthogonal frequency division multiplexing. Briefly talked about spatial multiplexing. This technology of spatial multiplexing is relatively new. How can you not use different frequencies and not different time and still get different steam of data? Is it the same as transmit diversity and receiver diversity? The underlying technique is the same but overall they are different. Whatever is considered not to be feasible 10 year ago is now feasible these days.

Receive Diversity:
Suppose you have a transmitter and you have one antenna, the transmitter T1 can send signal x at a time, and then you have one receiver but unlike the transmitter, you have diversity, you have multiple paths and not a signal path,
You can use two antennas (parallel antennas) at slightly different locations. Because the channels have multiple paths, they won't get the same signals. But R1 will get y1 and R2 will get y2.



H1,H2 is a complex number that is coefficient/gain and it is dependent on the degree of fading(depending on the type of obstructions) X can be a bit, but y is always a signal due to the presence of the noise component..
N1 and n2 represent Gaussian noise.

$y_1 = h_1 x + n_1$,
where h_1, h_2 represents a complex number, and is dependent on the degree of fading

$y_2 = h_2 x + n_2$,
n_1 and n_2 represent zero-mean Gaussian noise.

This n1, n2 are Gaussian noise and one of the properties of this gaussian noise is that they are not correlated. If you do the dot product, then E[n1.n2]=0

The receiver knows that it is getting only one symbol. It is observing two versions of the same symbol.
Because they know it is the same symbol that is coming, there is know that that symbol should come from y1 and y2 respectively. So that symbol is actually that the recipient will observe and is equal to the linear combination of distance.
And w1 and w2 are weights given to each antenna

$y' = w_1 y_1 + w_2 y_2$

Actual symbol that the recipient will observe

w_1 and w_2 are the weights given to each antenna.

What value of w1,w2 should be used so that y' has the highest amount of signal to noise ratio?
Here comes the mathematics of diversity.

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} h_1 x + n_1 \\ h_2 x + n_2 \end{bmatrix}$$

$$= \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} h_1 x \\ h_2 x \end{bmatrix} + \begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$$

W1H1+**w1H2**+**w2H1**+w2H2
= w1H1+w2H2
And we have chosen w1 and w2 as 1/sqrt(2)

Effectively if you look at this, this is actually the signal component on the receiver end and the other one is the noise component.

$$= \underbrace{\vec{w}^T \tilde{h} x}_{\text{Signal component on the receiver side}} + \underbrace{\vec{\tilde{w}}^T \tilde{n}}_{\text{Noise component}}$$

If your transmit power is P then

$$SNR = \frac{|w^T \tilde{h}|^2 P}{E\left[|\tilde{w}^T \tilde{n}|^2\right]}$$

W is a vector, h is also a vector. Square because power is always in terms of the square. Power is 0.5A^2.
Why do we have expectations?
N is Gaussian noise with a 0 mean.
If you look at the expected power, then the sum of the power will 0. That is why we have square.

Y is the summation of X(power of a signal) and N→noise power)(two different components)
We want to have the highest SNR, so we want to minimise the denominator.
We want to maximise the numerator and minimise the denominator.

$$\text{Maximize } |\tilde{w}^T \tilde{h}|^2 = (\tilde{w} \cdot \tilde{h})^2$$

$$= |\tilde{w}||\tilde{h}| \cos\theta,$$

where θ is the angle between $\tilde{w}$ and $\tilde{h}$.

If theta = 0 , cos theta = 1
Then effectively we need to choose w in such a way that has the same orientation as the channel vector. If you look at this

It has the same

$$|\tilde{w}^T \tilde{n}|^2 = \left(\begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}\right)^2$$

If you do the dot product then you get

$$= (w_1 n_1 + w_2 n_2) \cdot (w_1 n_1 + w_2 n_2) = w_1^2 n_1^2 + w_2^2 n_2^2 + w_1 \cdot w_2 \cdot n_1 \cdot n_2 + w_2 \cdot w_1 n_2 \cdot n_1$$

$$= w_1^2 n_1^2 + w_2^2 n_2^2$$

The reason these terms are 0, noise is always uncorrelated, when we do their dot product.
If you assume that Gaussian noise has a variance then the power of the noise remains the same irrespective of the frequency. So we can do like W1 and W2 are in our hand, we chose them as 1/sqrt(2). That is the value for with 45 degrees, we highest.

$$= \sigma^2 (w_1^2 + w_2^2)$$

If we assume that n1 and n2 are identically Gaussian Distribution Random variables with sigma square variance, the power of gaussian wave is given by variance.

$$SNR = \frac{|w^T \tilde{h}|^2 P}{E\left[|\tilde{w}^T \tilde{n}|^2\right]} \qquad SNR = \frac{P}{\sigma^2}\left(|h_1|^2 + |h_2|^2\right)$$

If you didn't have anything more than one receiving antenna, then you would have got p/sigma square.

Suppose you had only one receiving antenna.
Then, the $SNR = \dfrac{|h_1|^2 \cdot P}{\sigma^2}$

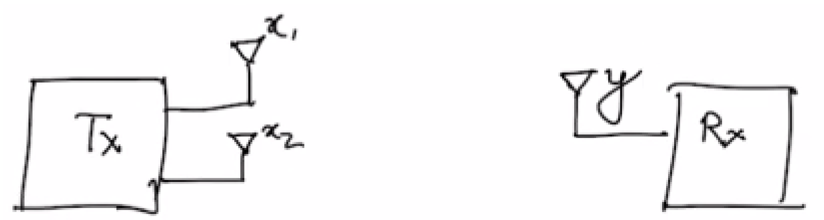
B = log(SNR)
Using this type of technique we are able to increase the SNR of the receiving signal. And this increase and SNR is increasing

Tradeoff: $\rightarrow$ Benefit: $\rightarrow$ SNR is increasing at no additional cost of bandwidth or transmission power.

Based on shannon capacity, we see, if we increase the transmission power, the SNR increases, but we are able to see without even increasing the power the SNR is increasing,

Loss: We now need more complex hardware at the receiving end. This is the idea of receive diversity. Why would hardware will increase? Using two antennas, the SNR is more It is not a major factor in practice. Can we actually do the same at the transmission side also? It is called transmit diversity. We can have more antennas on the transmission side. And we can also have one antenna at the reciver side.

Transmit Diversity



Case 2: multiple antenna at transmitter end.
We are receiving one signal.
The moral is we can actually identify what is being transmitted, at the transmitter side, we have to apply w1 and w2 at the transmitter end, that we were applying at the receiver side earlier. The receiver can infer from the data that was coming, but the transmitter side cannot infer easily, that info has to be received from the receiving side by the transmitting side. It is not really in the receiver end about adding/deleting coefficient. The receiver only has one symbol

Y = h1x1+h2x2+n
It is sum of multiple components of the signal.
H1 is the channel gain/coefficient of fading between T1 and R.
H2 is the channel gain/ coefficient of fading between T2 and R.
Now, what can be done to ensure that on the receiver side you are able to get the same type of SNR
The trick is you need to encode X1 and X2 as a function of x in such a way that SNR increases from the receiving end.

$$\tilde{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Y = h1x1+h2x2+n

$$y = h_1 x_1 + h_2 x_2 + n \qquad = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + n$$

If you want this to actually ensure that this comes to the summation of the power. You encode this X1 as conjugate of x

Encode x_1 as $h_1^* x$,
and x_2 as $h_2^* x$

Where h1* is the conjugate of h1
And h2* is the conjugate of h2.

The moment you do that,

$$\begin{aligned} y &= h_1 x_1 + h_2 x_2 + n \\ &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + n \\ &= \begin{bmatrix} h_1^* x & h_2^* x \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + n \end{aligned} \qquad = h_1 h_1^* x + h_1^* h_2 x + h_2^* h x + h_2 h_2^* x + n.$$

Now, these two are conjugates, a+bj and a-bj are conjugates of each other.
When you multiply them you get it as

$$(a+bj)(a-bj) = a^2-(bj)^2 = a^2+b^2$$

$$= x(|h_1|^2 + |h_2|^2) + n = |h_1|^2 x + |h_2|^2 x + x(h_1^* h_2 + h_2^* h_1)$$

$$= x(|h_1|^2 + |h_2|^2) + n$$

Here it is much more difficult for the transmitter, as he needs to know what is the channel condition being observed by the receiver side. Effectively this information/channel information has to be known at the transmission time.
X is just a bit, it is either 1 or 0

How it is similar to the earlier case.
In an earlier case, there was also the weight associated with the noise.
In this case, also noise is there.
Variance is sigma square so the power that we get

$$SNR = \frac{P(|h_1|^2 + |h_2|^2)}{\sigma^2}$$

Actually, the way to do it, we talked about the angle of w.
|w1| = 1/sqrt(2)

$$|w_1| = \frac{1}{\sqrt{2}}$$

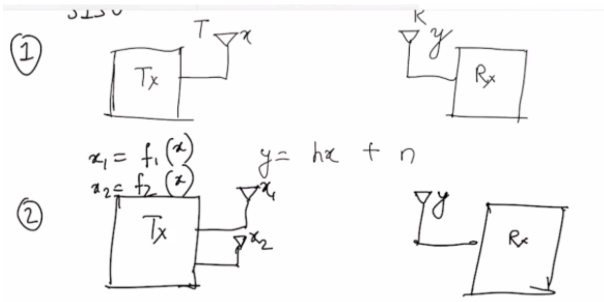
$$|w_2| = \frac{1}{\sqrt{2}}$$

Only in this case, this sigma square will come in the denominator. Otherwise, the w1 and w2 squares will be there.
Does this imply that the angle should be 45 between them? In the older days, lots of users said phones are manufactured by a lot of users. We cannot assume that all phones have multiple receivers. Even if the phone does not have multiple antennas, the transmission side. Is it possible that we avoid this? This is a very big challenge. In cellular networks, we get this behaviour. The channel information has to be same as the transmitter for it to be used. Control data is nothing but the channel state information. There is one way of stating it. This technique is called is still used by it. The wifi has access points. We will see at least 3 antennas. This is what we called as transmit diversity In that case, there is a type of encoding of data to sidestep this CSI. The way that can be done is you don't send one bit at a time as we encode one bit in two.Instead of sending only x, you send x' and x^2. Consider it as a signal. If you consider x as a signal then you are having this conjugate, x now has a conjugate x1 and x2. You are sending this x, now consider x as a signal. You can split this. Now x^1 becomes x1 and x2

$$x^1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \quad x^2 = \begin{bmatrix} -x_2^* \\ x_1^* \end{bmatrix}$$

Idea: Transmitting side needs to get information regarding the channel, which is hard in practice, but there is a way to avoid the information. The receiver side can infer the channel condition, but the transmitter faces difficulty.
How can we manage?
Idea is to handle this problem, the special type of coding, each time, if we have to send some bit x, we send it as x1 and x2. Whenever we are sending two bits, at that time you are sending one bit only. We encode the bits in such a way for the receiver to infer by some mathematical operations even though we don't know the channel's condition on the Tx side. We are losing the bandwidth, we are losing SNR by linear units it is ½

In the first bit, send x1 and in the next bit, you send x2. It is possible for the receiver to identify. There is a tradeoff. For the same amount of power, you are sending half the number of bits. It reduced channel utilisation. SNR falls by a factor of 2. In log terms, it is -3dB. log2=0.3010 You no longer need the CSI data. It is a headache to get the CSI data.
Maybe if the channel gain is reduced by 3db, we can use amplification by some medium
You have a single transmit antenna T and a single receive antenna and the transmit antenna sends X and R receive X.
y=hx+n, Now



$$y = h_1 x_1 + h_2 x_2 + n$$

We already have some formula of x_1 and x_2 as functions of both h_1 and h_2.

You have to access it.

$$y_1 = h_1 x_1^1 + h_2 x_2^1 + n$$
$$y_2 = h_1 x_1^2 + h_2 x_2^2 + n$$

If you do dot product between them,

$$x^1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad x^2 = \begin{bmatrix} -x_2^* \\ \end{bmatrix}$$
$$y_1 = h_1 x_1^1 + h_2 x_2^1 + n$$
$$y_2 = h_1 x_1^2 + h_2 x_2^2 + n$$
$$\vec{y_1} \cdot \vec{y_2} =$$

$$y_1 = h_1 x_1^1 + h_2 x_2^1 + n$$
$$y_2 = h_1 x_1^2 + h_2 x_2^2 + n$$
$$\vec{y_1} \cdot \vec{y_2} = \left(h_1 x_1 + h_2 (-x_2^*) + n\right)$$

$$\left(h_1 x_2 + h_2 x_1^* + n\right)$$

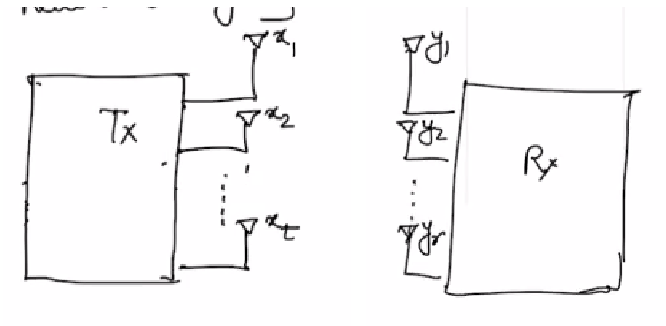
$$= \left(|h_1|^2 x_1 x_2 + h_2 h_1 (x_2)(-x_2^*) + h_1 h_2 x_1 x_1^* + |h_2|^2 (-x_2^*)(x_1^*) + n[\;\cdots\;]\right) = h_1 h_2 \left[x_1 x_1^* - x_2 x_2^*\right] + \left[|h_1|^2 x_1 x_2 - |h_2|^2 x_1 x_2\right]$$

These negative terms are reducing the SNR.
The negative terms would reduce the SNR values.
But we are effectively able to utilise the transmit diversity without knowing the value of h1 and h2.

There is a tradeoff and depending on the situation you are using it, it may or may not be a good thing.
If you have no way of sending the channel state information.

Can we actually utilise both?
We saw transmit diversity and receiver's diversity.
The question is can we use both?
How can we do it if we use both?
The answer is you can.

Suppose you have something like this.



$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix}$$

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = \begin{bmatrix} h_{11}x_1 + h_{12}x_2 + \dots + h_{1t}x_t \\ \vdots \\ h_{r1}x_1 + h_{r2}x_2 + \dots + h_{rt}x_t \end{bmatrix}$$

$$= \begin{bmatrix} h_{11} + h_{12} + \dots + h_{1t} \\ \vdots \\ h_{r1} + h_{r2} + \dots + h_{rt} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_t \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_t \end{bmatrix}$$

If we represent the channel as the matrix and

$$= \vec{H}\vec{x} + \vec{n}$$

Your job is the identify the value of this X matrix to decode the symbol being sent.

$$\tilde{x} = \vec{H}^{-1}(\vec{y} - \vec{n})$$

If you ignore the noise, then that will be somewhat equal to

$$\approx H^{-1}\vec{y}$$

You cannot ignore the noise.
Suppose you have more receiving antennas than the transmitting antennas, at this point, you get what is called a regression problem.(Linear regression where we extrapolate the curve to predict Y for some new X.)

You are now trying to fit some noisy data and

The above becomes a regression problem, and we want to solve it using the least squares method.

$$\tilde{x} = \underbrace{(H^T H)^{-1} H^T}_{\text{Formula comes from the way regression works}} \tilde{y}$$

We are using zero-forcing receivers because we are trying to minimise the noise.
This is the way we are actually able to utilise the total power of this.
The total power is the overall summation.
Total of H12,..., H1T,...,HR1,...., Hrt
And the total SNR

Zero-forcing receiver; because we are trying to minimize the noise.

$$H_{12}, \ldots, H_{1t}, \ldots, H_{r1}, \ldots, H_{rt}$$

$$SNR = \frac{P}{\sigma^2}\left(\sum_{i=1}^{r}\sum_{j=1}^{t} \|H_{ij}\|^2\right)$$

This is the total amount of gain, it is known as total diversity gain. Because this gain is coming from both transmitting and receiving diversity.
That is why this is the diversity gain.
If we assume that all the w values are 1/root 2, then the total diversity gain is given by the product of the number of receive antennas and the number of transmit antennas.

Lecture 6

- Spatial Multiplexing
- Multi-user MIMO
- Massive MIMO

Continuation of MIMO
We have multiple antennas on the receiver's side and a single antenna on the transmitting side. It is possible to enhance the bandwidth using this technique. Next, we discussed that there are multiple antennas on the transmitter side and a single antenna on the receiver side. The SNR has the same formula.

There is something called receiver diversity gain where you can use multiple receiving antennas, to sum up, the SNR. You have something called transmitting diversity gain where you can use multiple transmitting antennas, to sum up, the SNR. It required CSI data on the transmission side, it can be done at the sidestepped using some coding scheme at the cost of SNR. How can we combine this receiver diversity gain and the transmitter diversity gain?

Requires another type of formula at the receiver end to combine/sum up the SNR.
It was all about diversity gain. SNR was the same in the previous case but there was the catch, we needed the Channel state information(CSI) on the transmitter side. Now, we were able to avoid it by using some spatial encoding but that came at the cost of some amount of efficiency. We showed with mathematical analysis too. On the receiver side, we can get it by multiplying the response of the channel and we can retrieve X using the receiver. We can use a form of linear regression whose main aim is to reduce the noise.

$$\tilde{x} = (H^T H)^{-1} H^T \tilde{y}$$

$$\tilde{y} = \tilde{H}\tilde{x} + \tilde{n}$$

$\tilde{H}$ is the channel matrix
$\tilde{x}$ is the bit vector transmitted
$\tilde{n}$ is the vector of Gaussian noise
$\tilde{y}$ is the vector of received signal

$\tilde{x} \rightarrow m \times 1$ vector
$\tilde{y} \rightarrow n \times 1$ vector
$\tilde{H} \rightarrow n \times m$ matrix $\rightarrow$ complex matrix
$(m \times n)$

Suppose you have **m** number of antennas on the transmitter side and **n** number of antennas on the receiving side.

We have already explained that this is a complex matrix. Now, many of you might have learnt in the linear algebra course that any complex matrix can actually be factorised into the following form.

$U \Sigma V^*$; $U = n \times n$ matrix
$\Sigma = n \times m$ matrix
$V^* = m \times m$ matrix
[Singular value decomposition] ;

$\Sigma = n \times m$ matrix
$V^* = m \times m$ matrix
[Singular value decomposition] ; U and V are the diagonal matrices

On the transmit side, send $(\tilde{V}\tilde{x})$.
On the receive side, apply $(U\Sigma)$; $\Sigma = [\delta_1, \ldots, \delta_t]$

$y_1 = \delta_1 \tilde{x}_1 + n_1$ $\quad U\Sigma\tilde{V}x$
$\vdots \quad \vdots$

U and V are the diagonal matrices. The idea is if you send Vx and on the transmitter side, apply $U\Sigma$ on the receiver side.

$$y_1 = b_1 \tilde{x}_1 + n_1$$
$$\vdots$$
$$y_t = b_t \tilde{x}_t + n_t$$

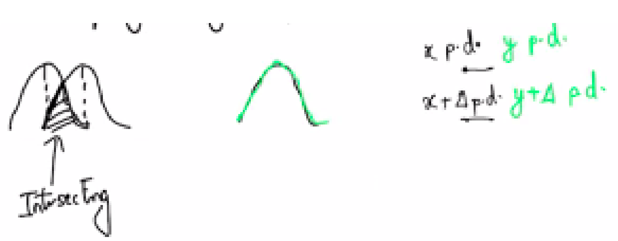
If the sigma values are all distinct then you can solve the equations to get the distinct values of the vectors, y1,y2,...,yt.
Using this, you are effectively instead of getting the same values, you are able to recover the values of x1,x2,....,xt.
So you are able to get different bits of information but over the same channel.

Time-division frequency/frequency division multiplexing. If you have lots of paths then you have different values of sigma1, sigma 2 and so on.
This is called spatial multiplexing.
If any of the sigmas is the same, then we don't have any solution.
We are sending two signals on the same channel, still, it will be able to decode without any interference.
There are two types of gain, one is called beamforming or diversity gain.

Spatial Multiplexing

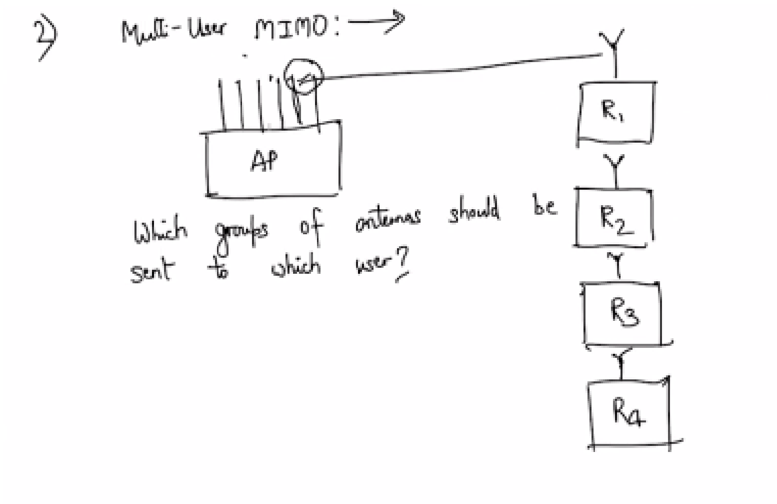
1) Beamforming or diversity gain = $\underset{\substack{\uparrow \\ \text{no of receive antennas}}}{r} \times \underset{\substack{\uparrow \\ \text{no of Transmit antennas}}}{t}$

R is the number of receiving antennas and T is the number of transmit antennas. R*T is the gain in the SNR.
There is something called multiplexing gain or gain in the number of bits you send which is also called min(r,t) but also is a property of the channel.
If there is a line of sight condition then the main components may go directly with scattering then spatial multiplexing might fail.
In practice, we don't use a lot of antennas.
Channel state information is needed on the receiver side, you not only wanna increase SNR but also you want to send multiple bits of information.
MIMO is all three i.e. CSI, SIMO, MISO
The key difference between MIMO and Transmit beamforming, Receive Beamforming. We wanted to increase SNR using multiple antennas, Today the goal is to use the same antennas I want to send multiple bits in parallel. On the same channel. Channel can be multiple things. At the same time and at the same frequency to send the information. If the channel has the property that there are lots of problems, then there are lots of interference, so that the receiving side can decode the signal.
Start small antennas and then keep on increasing.
A channel means at the same time and same frequency.

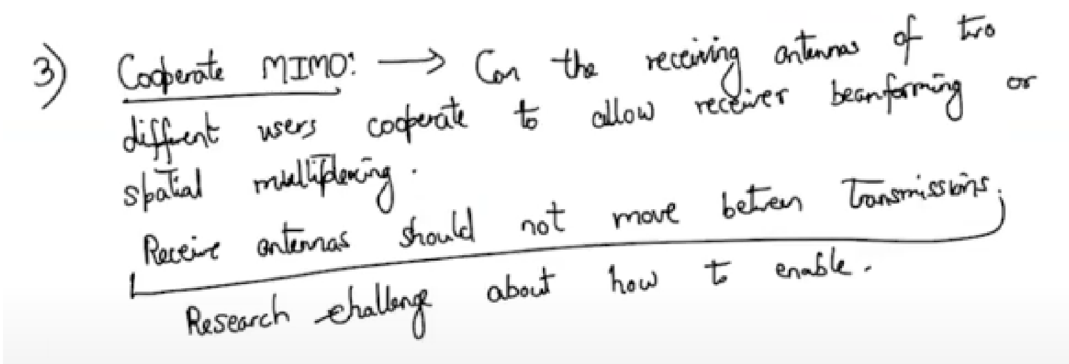


Delta depends on the channel as the antenna is not moving.
The most intense mathematical derivations are finished. These extensions are actually what we are trying to explore. That technique is called single-user MIMO/ point to point MIMO.

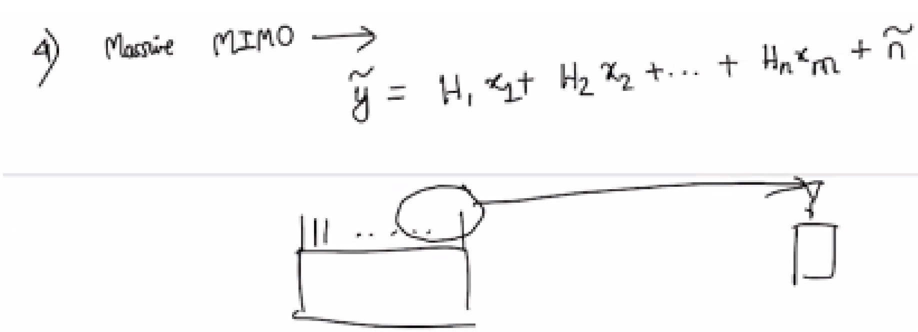
In this, only one Tx and one Rx is used. We can also use a Multi-User MIMO.



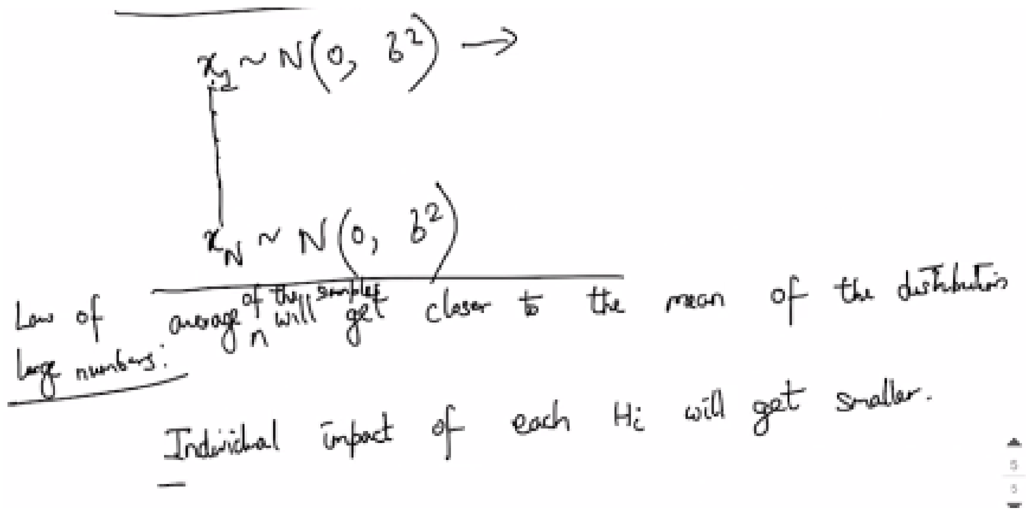
Sometimes we use transmission time in parallel and for others, we can use non-parallel. We can also do some sort of schedule based on channel-based properties.
This type of Multi-User MIMO is used in cellular networks.



The third extension is cooperative MIMO.
Among the receivers can the receiving antennas of two different users themselves collaborate to allow receiver beamforming or spatial multiplexing. The challenge is the receive antennas should not move between transmissions.
We have one antenna at the transmitter side and many antennas at the receiver side.
The idea is called a massive MIMO. If you keep on increasing the antennas, suppose you have 64 antennas instead of 3, then what you can do, on the receiver side, you can use 16 antennas to send data to one of the recipients, in that case, one property of the channel. It will get summed up.



Because, this h1, h2 and all these are almost from the same location, they have roughly the same distribution but the same values, in expectations, these have the same value. A property that, gradually as the number of transmit antennas increases, the summation of the H values would tend closer to the mean from the law of large numbers.



Suppose you are taking multiple channels, suppose you have variance sigma square. Suppose you take one number, what do you expect that number to be, we don't have any idea, it depends on the sigma.
If you keep taking on these numbers gradually, you will see that gradually as the number of samples you are taking, then the average of the samples will be the same as the mean of the distribution, which is equal to 0 in this case. In the same way, when you are using a lot of antennas, when we have a channel impulse variable, the gradual response will become closer to the mean. In the law of large numbers, it can be from any distribution, because binomial tends to be gaussian.
This is known as channel hardening. You no longer need CSI data as often. And this concept of massive MIMO is used in the upcoming version of 5G.
Advantages of Massive Mimo
More Bandwidth as carrier frequency is so high.
Multipath Diversity goes down/decreases.
How can the CSI data be sent to the transmitting side if required?
Time-division duplexing→ If the protocol says that periodically let the receiver send some recognised signal, so the transmitter side can understand the CSI by reading what signal it gets.
Frequency division duplexing→ The Tx sends the recognised signal, so the receiving side first infers the CSI and then sends it back to the transmitting side. It depends on the protocol, both of them have advantages and disadvantages.

FDD disadvantage→ Require a specific channel to send back the CSI to the Tx. This is used in LTE networks. On the other side.
TDD Disadvantage→ Tx cannot all the time understood from the receiver side, if the receiver's transmission signals may not be calibrated as receivers may have diverse power properties of the receiver. If you look at the electronic property then you will see that all Tx and Rx have their own sort of hardware imperfection. Because of these hardware imperfections, they are requested to send at the same power and frequency, they are a significant amount of difference in terms of power.
Same phones, still the power transmission by two phones have different unless the manufacturer has done special thing to avoid it.
Why is there a difference in the Power?
Suppose we have a receiver R1 and R2, we assume that they are identical. Whenever there is an analog component in any hardware, then they can never be identical in practice. And all antennas have analogue components.
Exact signal signature before it is converted in digital form, then you can identify which hardware is generating it.

1. Run a control system over a wireless network

Control system over a wireless network

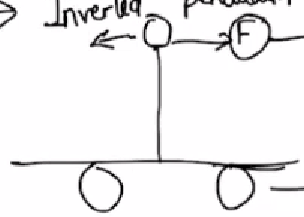
Raspberry Pi / Raspberry Pi Zero and check ho

closely you can control.

→ Inverted pendulum on a cart.

2.

→ Inverted pendulum on a cart.



→ A contolled force ~~by~~ to be infrred by calculation by a control system

⇓

controller which is getting

the position and velocity of the pendulum as input and computing f as output.

① Investigate the impact of increase in latency.

the position and velocity of the pendulum as input and computing f as output.

② How does the variance in latency affect?

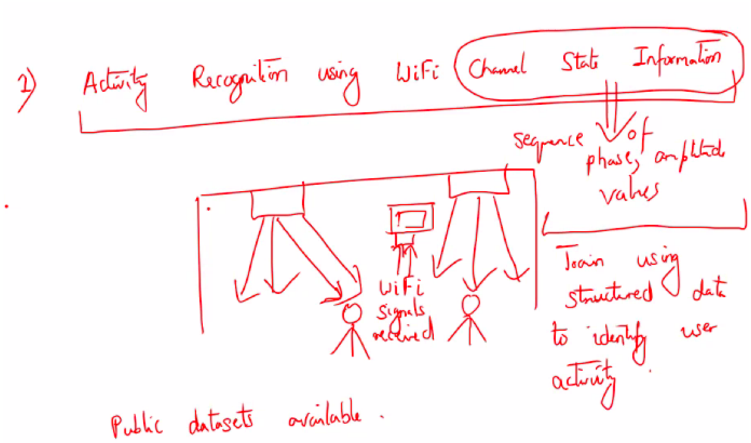
③ Can you use a more sophisticated co

Lecture 7:

- Activity Recognition using WiFi Channel State Information Data
- Analysis of app response time on smartphones – add delay to the network, and see how that affects the app response time
- Running Android on personal computer and checking its performance
- Simulation of video streaming over different network patterns
- Creating a closed loop controller that runs over wireless network
- Fingerprint-based localization of users
- Study of the impact of mobility on performance of wireless networks
- Localization of aircrafts/drones using RSSI and TDoA
- Study of TCP variants
- Running multipath TCP on either smartphones or Raspberry Pi's and studying its performance

Activity recognition using Wifi Channel State information Data

Whenever wifi access points are deployed on the wall, then your signals are propagating.
If you have some sort of sense, on a smartphone, which has a wifi device, then how these wifi signals are changing here, then we can tell ow the user is moving and what kind of activity the user is involved in. This information is known as channel state information.
It consists of sequence of some phase and amp values.
If you are able to train this using some structured data, then you can use this to identify the user activity.



2. Analysis of app response time on smartphones

We have a technique of measuring an app response time. Suppose I press post Reel on instagram, then it is known as app action. It takes some time to get the response time on the UI, then usually users love to see faster response. Using the system UI automation, which identifies the response time of the app.

Ai Automator → Measuring the app response time.

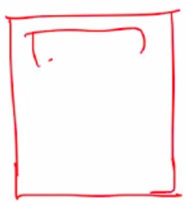
We always love faster response time.
How much does adding network delay or packet collision hurt the response time

2) Analysis of App Response Times on Smartphones

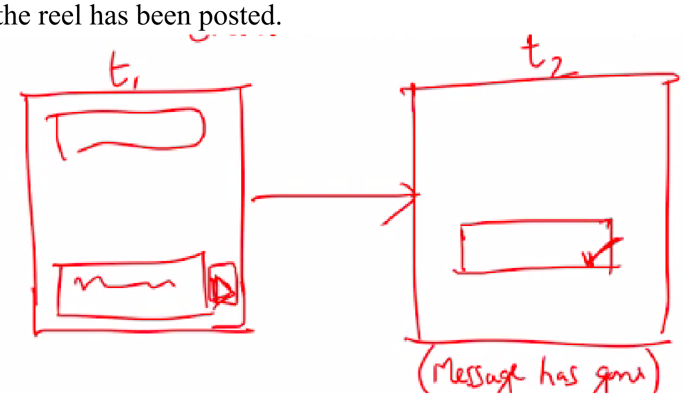
→ We have a technique of measuring the app response time.

→ Our AI Automator measures the app response times.

→ How much does adding network delay or packet collisions hurt the response time?



You type something here and press the post Reel o instagram, and this message show this tickmark, which implies that the reel has been posted.



$RT = t_2 - t_1$

Response time is t2-t1. User did some action and the action gave some response time, the smaller RT is considered to be better, which means a faster response.

As you are sending the data to the server if you are in a place which is far from a major city.
Fading effects are handled well.

In practice, when you do some amount of measurement, then you will do the experiment in good network condition.
You have to run the Android system on a personal computer, then check how it turns out in app.

3. Running android on personal computer and checking it's performance

If you have a better phone, then you should see a faster response. There is a limit to how fast a response we get. Smartphones response times are limited by the network.
You do not know how fast the response can be in practice.

That is why we can run an automation system on a desktop and see what is the fastest possible response time.

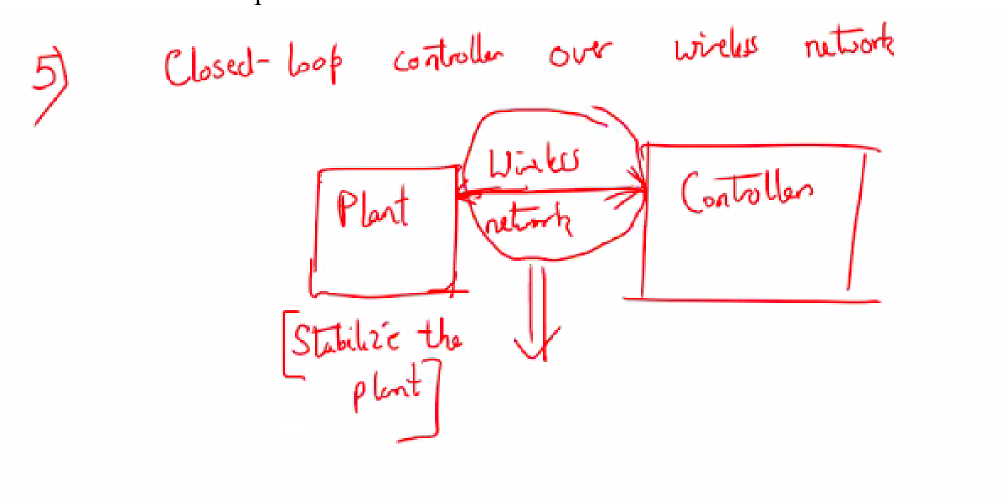
4. Simulation of video streaming over different network bandwidth patterns.

In practice, there are a lot of BW traces, whenever we want to do the video streaming, in some circumstances.
The reason is wireless network bandwidth which does this.
Open source simulators of video streaming systems can be used to combine the bandwidth traces to understand why video stops in practice.
This can give you an idea what can an actual challenge wireless channel faces

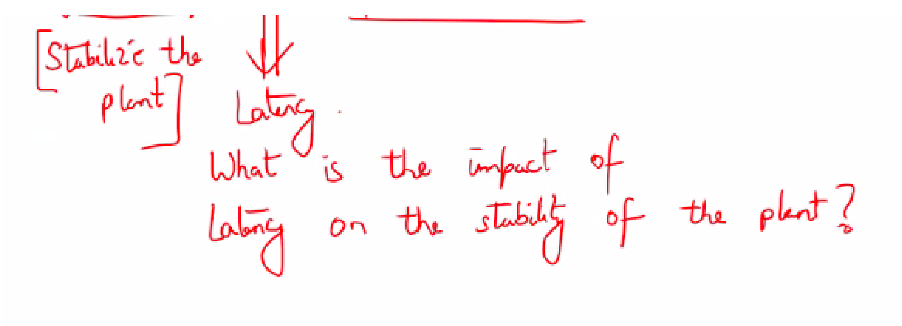
Simulation on fading condition.

5. Closed-loop controller over a wireless network

The idea is we have to have some kind of control system and you want to control something and you have to control a drone and anything that you need to control. It is known as a plant, and the controller is attached to this, in practice, we don't want the controller to be attached because that has to be relatively less computationally expensive or it became very expensive. In the factories, they are trying to bring a wireless network here and here you have a controller.
The problem: That can control this plant based on some data. It is sending the data, which is specifying to take some action to stabilise the plant.



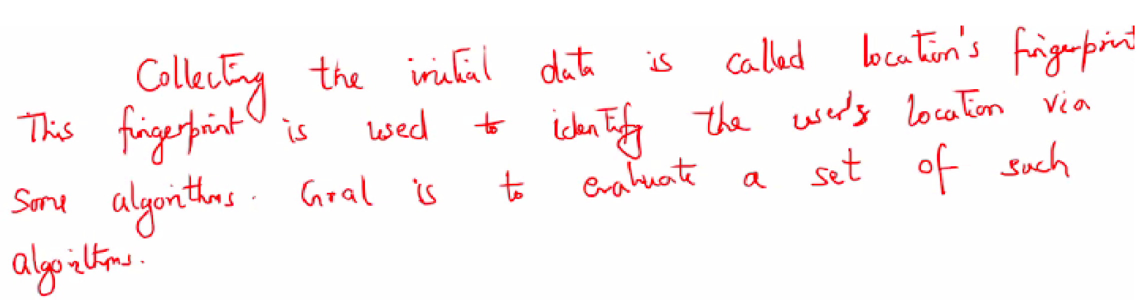
You will start to see latency, what is the impact of latency on this.



6. Fingerprint-based localisation (**Indoor**), Outdoor also

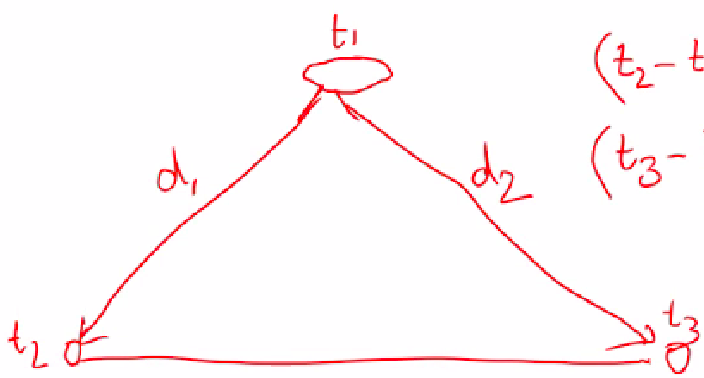
It is like in practice, technically, if you did not watch this fading and blocking, you could have used the power received(Received signal strength) should be enough to identify the location of the user-provided the location of the transmitter is enough.
How how can you identify the location of the user?

The technique is you have to do something similar to activity recognition. You have to study what is the power received there and then go to the power received and go there and you have to match with the previous data.
This is actually a collection of initial data is called a location's fingerprint.



Localization of aircrafts/drones using RSSI and $\underbrace{\text{TDoA}}_{\text{Time difference of Arrival}}$

These sensors will receive the signal at a different frequency.



$$(t_2 - t_1) \propto d_1$$
$$(t_3 - t_1) \propto d_2$$

Time difference of arrival of signals to localise the aircraft control.
Dataset also available on this.
We need to study how time difference of arrival works.

Study of TCP variants

TCP → Basic TCP protocol used
TCP has variants to adapt to either wireless networks or even to ethernet.

How you change, directly effect TCP performance. If you send in wireless networks how these things are actually effecting the performance using web streaming, browsing or any of the one application with multiple variants of TCP.

Study of TCP Variants

TCP → Basic TCP protocol used.

TCP has variants to adapt to either wireless networks or to data center networks or even to Ethernet.

Any one application with multiple variants of TCP.

Video → Performance in terms of resolution / rebuffering.

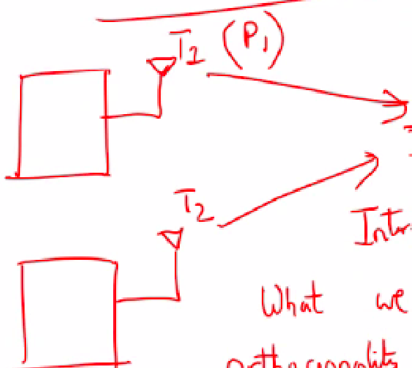
TCP taho and reno and CUBIC.
These have an effect on the performance of the application.

Successive Inference Cancellation:

In practice, suppose you have two transmitters Tx1 and Tx2, if they both send power together, then interference will happen. This was considered to be unwanted as interference makes hard to decipher signals.

We used the principle of orthogonality.
Using OFDM, TDM was two technique, spacial multiplexing was another technique, all of them used orthogonalituy

Successive Interference Cancellation



Interference (Unwanted)

Interfrence makes it harder to decipher signals.

What we did was to use the principle of orthogonality in various ways.
Using OFDM, TDM, spatial multiplexing.

Is this the best we can do?

$$C_1 = B \log_2\left(1 + \frac{P_1}{N_0}\right) \qquad C_2 = B \log_2\left(1 + \frac{P_2}{N_0}\right)$$

What if we checked P1 and P2 at the same time, then we would not be able to identify both of the signals. Which is that suppose we send signals from T1 with power P1 which is greater than P2. It has to be significantly greater. P1 >> P2

Because, this power is actually, the way the recipient twill see the signal it will perceive that it will be added.

$$B \log_2 \left(1 + \overset{I}{SNR}\right) = B \log_2 \left(1 + \frac{P_1}{P_2 + N_0}\right)$$

C1 will be the capacity of the signal.
Because P1 is significantly greater than P2 + N0
This will have some value greater than 1
Effectively this will have a large value, then indicates that the T1's message can be decoded.

Even if there is an interference signal, you are able to detect T1's message. Now, if you actually know what is actually T1's message, whatever signal you have received, you can subtract whatever you have received(ideal signal decoded by the recipient i.e. the signal of T1) then you get the residual power as the transmitted signal of T2.

If you have some signal strength that allows you a sufficient amount of information using Shannon's capacity formula.If the signal strength is low, channel capacity is also low.

You have sufficient power to identify that I can decode it.
Mathematically, P1 > P2+N0 then log(2+alpha) then you get more capacity than the bandwidth allowed.

If you know what T1 sent, you know the modulation of T1.
Subtract T1's part, and let the residual power, i try to identify the signal sent by the other user.
What is the exact amount of capacity?

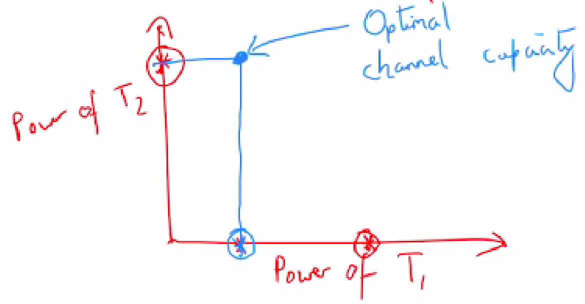
Effectively,

$$C_2 = B \log_2 \left(1 + \frac{P_2}{N_0 + \mathcal{E}}\right)$$; $\mathcal{E}$ is the additional noise introduced by the subtraction

C2 is transmission capacity in bits/sec
B is the channel bandwidth

Even though you have not used orthogonally, you have used non-orthogonal signals, you are able to detect.
Non-orthogonal signals can be detected.
This technique of identifying the signals is known as successive interference cancellation.(SIC)
It was proposed in 5G.
SIC gives new multiple access techniques that use SIC. It is called NOMA(Non-orthogonal multiple access)
If you do something like this, then this is the actual point of optimal channel capacity.

Non-orthogonal signals can be detected.
Successive Interference Cancellation (SIC); proposed in 5G; the new multiple access technique that uses SIC is called NOMA (non-orthogonal multiple access).

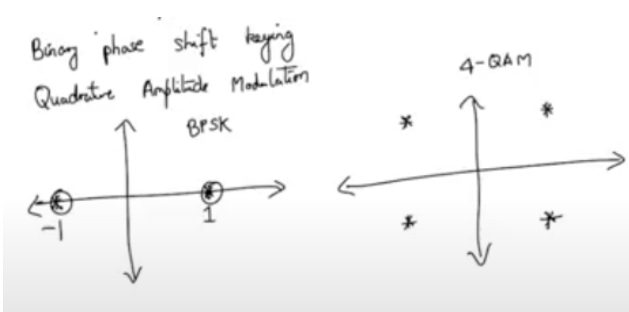


The receiver knows the modulation scheme of each of the transmissions.
This can be done when multiple transmissions are also there.

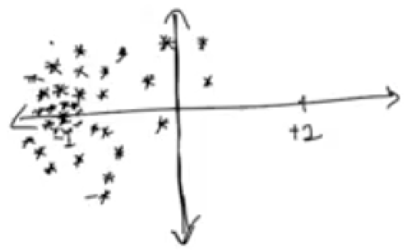
$$C_1 = B\log_2\left(1+\frac{P_1}{P_2+P_3+\dots+P_n+N_0}\right) \qquad P_1 > P_2 > P_3 > \dots > P_n$$

Sometimes you will find it too hard to decode.
Currently upto 4 transmitters in State of the art.
You need a series of rectangles.

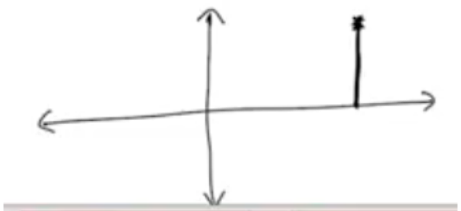
Bit error Rate (BER) and Simple Error Rate(SER)
When we discussed modulation, we discussed Binary phase shift keying or Quadrature Phase shift keying and we discussed Quadrature amplitude modulation. And if we represent it as a complex number then BPSK looks like



Only noise, that we have is additive white gaussian noise (AWGN). Effectively if we are sending some signal x, then we receive y(t) = x(t) + n(t).
Now, in general, if we are sending this data, and as the noise is stochastic in nature, then it crosses the boundary, it says noise is +1 and not -1 and there will be bit error.



In probabilistic terms, in terms of amplitude, we can say that we would have send the data like a fixed one, if there were no noise. It is white noise, it does not have any bias and a variance of sigma square.
N(A,sigma^2). We are applying a threshold. Any bit which is received which is greater than 0, If y(t)> tau, then we estimate x(t)=1. If it is less than tau, then x(t)=0.



If $y(t) > \tau$ then we estimate $x(t) = 1$
$\leq \tau$ then we estimate $x(t) = 0$.

Threshold

If you assume tat both 0's and 1's are equally likely than, we can show that the error is minimised if tau=0.

Probability that the data you are sending but your y(t) is less than tau
Or you are sending 0 but the data received is on the right side of the tau.

$$P_e = P(x(t)=1)\, P(y(t) \leq \tau \mid x(t)=1) + P(x(t)=0)\, P(y(t) > \tau \mid x(t)=0)$$

$$= \frac{1}{2} P(y(t) \leq 0 \mid x(t)=1) + P(y(t) > 0 \mid x(t)=0)$$

$$= P\left(Z(t) \geq \frac{1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right) = Q(\sqrt{SNR})$$

1 − CDF of std normal is called as the Q function

Power of signal = 1 (A^2)
Power of noise = σ^2

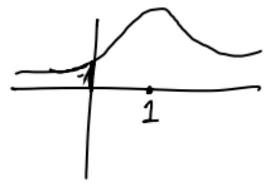
$$P_e = \frac{1}{2}T_1 + \frac{1}{2}T_2$$
$$= \frac{1}{2}Q(\sqrt{SNR}) + \frac{1}{2}Q(\sqrt{SNR})$$
$$= Q(\sqrt{SNR})$$

$$P_e = P(x(t)=1)\,P\big(y(t) \le \tau \mid x(t)=1\big) + P(x(t)=0)\,P\big(y(t) > \tau \mid x(t)=0\big)$$

$$= \frac{1}{2} P\big(y(t) \le 0 \mid x(t)=1\big) + \frac{1}{2} P\big(y(t) > 0 \mid x(t)=0\big)$$

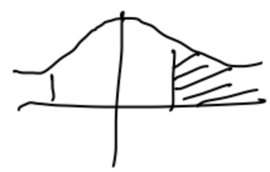
$$= \frac{1}{2} T_1 + \frac{1}{2} T_2$$

In T_1, $y(t) \sim N(1, \sigma^2)$



$$P\big(y(t) - 1 \le -1\big)$$

$$P\left(\frac{y(t)-1}{\sigma} \le \frac{-1}{\sigma}\right)$$



$$Z = \frac{y(t)-1}{\sigma} \sim N(0, 1)$$

$$P\left(Z(t) \le -\frac{1}{\sigma}\right)$$

$$= P\left(Z(t) \ge \frac{1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right) = Q\left(\sqrt{SNR}\right)$$

1 − CDF of std normal is called as the Q function

Power of signal = 1 (A^2)

Power of signal = 1 (A^2)

Power of noise = σ^2

$$T_2 = P\big(y(t) > 0 \mid x(t) = 0\big)$$

In T_2, $y(t) \sim N(-1, \sigma^2)$

$$P\left(y(t) > 0 \;\middle|\; y(t) \sim N(-1, \sigma^2)\right)$$

$$= P\left(\frac{y(t)+1}{\sigma} > \frac{1}{\sigma}\right)$$

$$\left[Z = \frac{y(t)+1}{\sigma} = \text{std normal distribution} \right]$$

$$= P\left(z > \frac{1}{\sigma}\right) = Q\left(\frac{1}{\sigma}\right) = Q\left(\sqrt{SNR}\right)$$

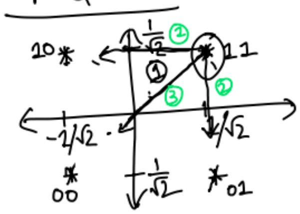
$$\boxed{Var(z) = \frac{1}{\sigma^2} Var(y(t)) + 0}$$

$$P_e = \frac{1}{2} T_1 + \frac{1}{2} T_2$$

$$= \frac{1}{2} Q\left(\sqrt{SNR}\right) + \frac{1}{2} Q\left(\sqrt{SNR}\right)$$

$$= Q\left(\sqrt{SNR}\right)$$

4-QAM



$$P_e = P\left(Re(y(t)) < 0 \;\middle|\; Re(y(t)) \sim N\left(\frac{1}{\sqrt{2}}, \sigma^2\right)\right)$$

$$= \frac{Re(y(t)) - \frac{1}{\sqrt{2}}}{\sigma} = Z$$

Grey code encoding

$\Downarrow$

Probability of losing both bits reduces

$$P\left(\frac{Re(y(t)) - \frac{1}{\sqrt{2}}}{\sigma} < \frac{-\frac{1}{\sqrt{2}}}{\sigma}\right)$$

$$= P\left(Z < \frac{-\frac{1}{\sqrt{2}}}{\sigma}\right) = P\left(Z > \frac{1}{\sqrt{2}\sigma}\right)$$

$$= Q\left(\frac{1}{\sqrt{2}\sigma}\right)$$

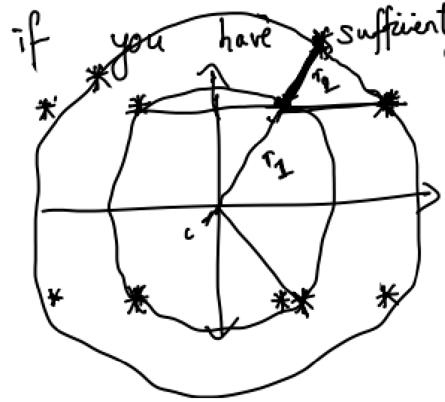
$$= Q\left(\frac{\sqrt{SNR}}{\sqrt{2}}\right)$$

$$\text{Total SER} = Q\left(\sqrt{SNR}\right) + 2Q\left(\frac{\sqrt{SNR}}{\sqrt{2\cancel{SNR}}}\right)$$

$$BER = 2Q\left(\frac{\sqrt{SNR}}{\sqrt{2\cancel{SNR}}}\right) + 2Q\left(\sqrt{SNR}\right)$$

{ If you pack more bits into the same signal, you are gaining in terms of data rate. But you are losing in bit error data.

You ~~one~~ should use higher degrees of modulation only if you have sufficiently high SNR.



8-QAM

How should I encode the symbols so that the bit error rate is minimized?

$$r_2 = 2r_1$$

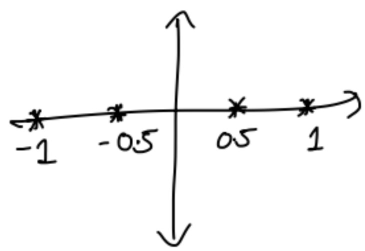
$$r_2 - r_1 = r_1$$

WiFi-6 (ax) → 1024-QAM, {BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM}

WiFi-5 (ac) →

LTE → 64-QAM

Ethernet → 4096-QAM



Why do we use 16-QAM and not 16-PSK?
But we use 4-PSK (QPSK) and not 4-QAM?

If you compare the formula's of BER,

$$BER(4\text{-QAM}) > BER(QPSK)$$

$$\text{But } BER(16\text{-QAM}) < BER(16\text{-PSK})$$

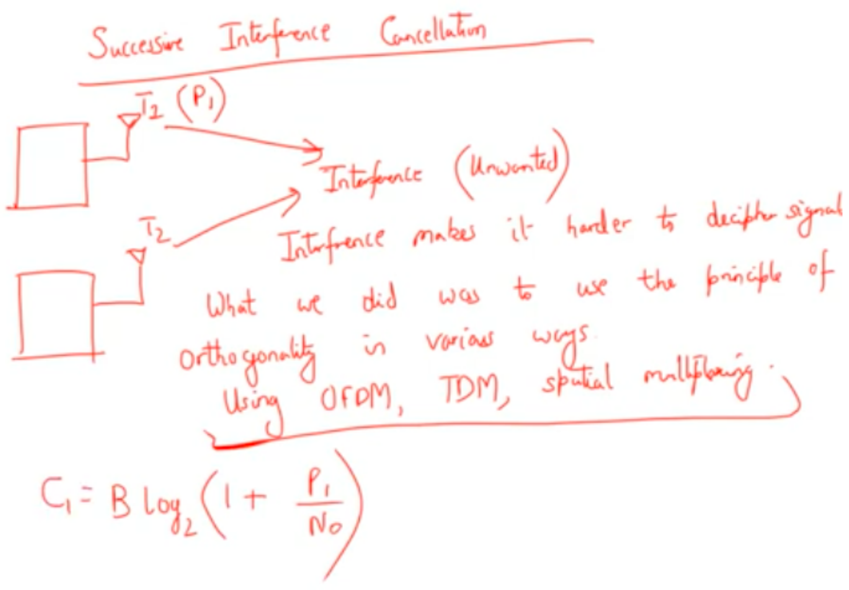
Lecture 8: Project Discussions

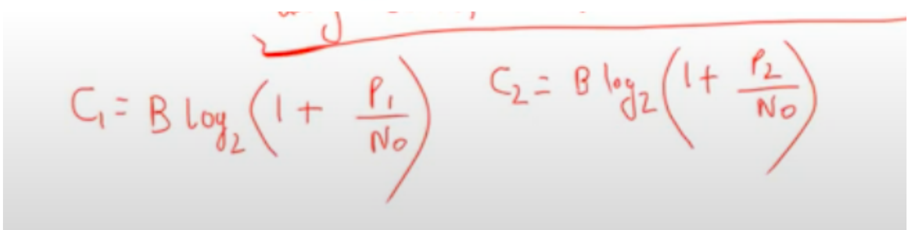
Successive Interference Cancellation

So Far, we have studied, in practice we have two transmitters T1 and T2 and if both send out power together, then interference occurs, then it is usually considered as an unwanted thing, as it makes harder for us to decipher signals. We used this principle of orthogonality in various ways. Using OFDM was one technique or avoiding interference, using TDM was another technique, the Spacial multiplexing was yet another technique, all these techniques used orthogonality.

If you are having this problem, then highest amount of power or bit rate is limited by shannon's limit.

If T1 has power P1 then you will get





Is this the best that we can do?

What if we had checked both P1 and P2 at the same time, then what would have happened?

There would be interference which means that you would not be able to identify any of the signals, but there is a way to solve that, which is that suppose we send a signal from T1 with power P1 which is significantly greater than P2. Now, because, receiver will see the signal as Blog(1+SNR) and other signal is added, so it will become $C1 = Blog_2(1 + S^iNR) = Blog_2(1 + \frac{P_1}{P_2+N_0})$

Because, P1>>P2+N0 then the value will be greater than 1, then this will be having a large value which indicates that the T1's message can be decoded. Even if there is an interference signal, technically you will be able to detect T1 message, but if you actually know what T1 message is, then whatever signal you have received, you can subtract whatever you have got from that signal, and then forward it also passing on what T2 has sent.

So if you subtract the ideal signal of T1 decoded by the recipient, then you get the residual power as the transmitted signal of T2. Superimposition happens with longitudinal waves. But you can see it as a superimposition in case of radio frequency waves.

When can a message be decoded?

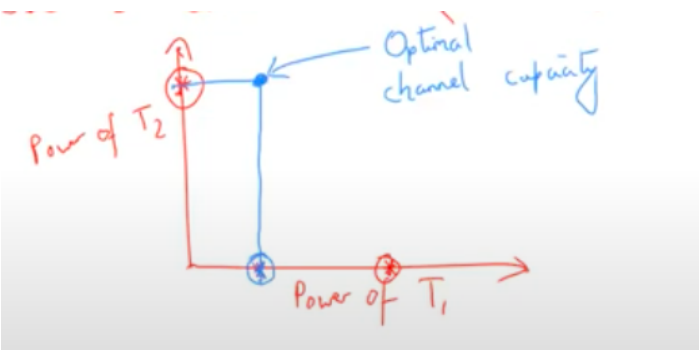
If you have some signal strength that allows you a sufficient amount of information, that can be measured using the Shannon's capacity formula whether/whatever the signal. If the signal strength is low, the channel capacity is also low. Because, P1 > P2, so effectively whatever signal strength you have is better for P1 than for anyone else. Then you have sufficient power to decode it.

Mathematically you can see that P1 >> P2+N0 and if you take log(2+alpha) is you have more capacity than what the bandwidth is allowing you.
But now if you know what T1 sent, then you know the modulation scheme of T1, then you subtract whatever T1 has got, and let me get the residual power, and in this power, try to identify the signal sent by another transmitter. Now, in this residual power, what is the exact amount of capacity. As it is sent at the same bandwidth, then effectively, I am assuming that the subtraction added some additional noise, then it can be written as epsilon.

B is the Bandwidth, and C is the transmission capacity in bits per second.

$$\underbrace{C_2}_{\text{Transmission capacity in bps}} = \underbrace{B}_{\text{The channel bandwidth}} \log_2\left(1 + \frac{P_2}{N_0 + \varepsilon}\right); \ \varepsilon \text{ is the additional noise introduced by the subtraction.}$$

If this power is again sufficient, then again you can decode.
Here you have not used orthogonality, you have used non-orthogonal signals. Non orthogonal signals can also be detected. Now hardware is getting better, so we no longer need orthogonal signals, this technique for identifying signals is called successive interference cancellation(SIC). It was proposed in 5G. Currently, no one has used it. We will see it in the next deployment of 5G problems.
The new multiple access technique that uses SIC is called non- multiplex.
It is used to improve bandwidth.



The receiver already knows the modulation scheme of each of the transmitters.
Because, we are subtracting the power, to get the residual power. It can be done for multiple signals. It is possible to generalize it for multiple signals. The receiver would know how many signals are been sent.

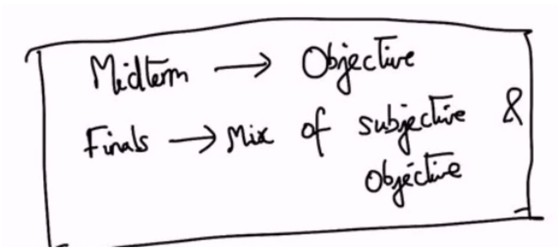
$$C_1 = B\log_2\left(1 + \frac{P_1}{P_2 + P_3 + \dots + P_n + N_0}\right)$$

$$P_1 > P_2 > P_3 > \dots > P_n$$

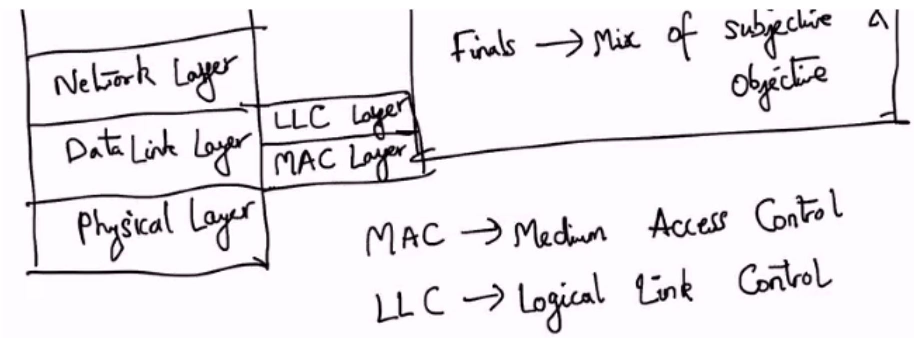
The limit depends on the quality of the channel as you are adding epsilon noise, then how much noise you are adding, that governs how much transmitter data we can send.
Currently, people have tried only 4 transmitters.
It is because of the noise. We need power. This is a rectangle. The more you can create asymmetric rectangles. It will be a hyperplane.

Lecture 9

- Concept of simulation, emulation and real experiments, with specific emphasis on evaluation of wireless network technologies
- Introduction to Data Link and MAC Layer
- Concept of ALOHA
- Concept of CSMA/CA and CSMA/CD



Communication networks are different. We didn't go into implementation, we know conceptually how it worked.
The physical layer is discussed in a good sense, We will get into the data link layer. Usually, Datalink is divided into two sub-layers, one is MAC which is Medium Access Control and the other one is the Logical Link Control(LLC) layer



Lots of research into MAC layers. There is one common thing in DDL and PL. These two layers are always implemented in hardware. You cannot change the physical layer technology because these are implemented in hardware. The Transport and Network layer are implemented in the OS code/Kernel Code. Above that, we have an application layer that is actually user-level code.

What is the easiest way to do the experiments, one is called simulation. NS3 is used for simulation. Simulation depends on assumptions and also on how do we model the environment. All models introduce some errors. Some errors will be introduced. When we say that we have this rayleigh fading, rice fading. These are models but also then you require some additional effort to ensure that these models are approximations of what has happened in reality. We have assumed that we have Gaussian noise. But it is not always true. It is sort of difficult to argue that simulation is enough to do everything. Conceptually in practice, we have three types
Emulation is closer to reality because it utilises the data from the real world and tries to imitate real-world actions.

EMULATION(example)
Suppose we want to find out in practice what is the behaviour of a video.
We are watching a video, we see the latency of the network. But in practice, if we move around, what will be the behaviour in the real world.

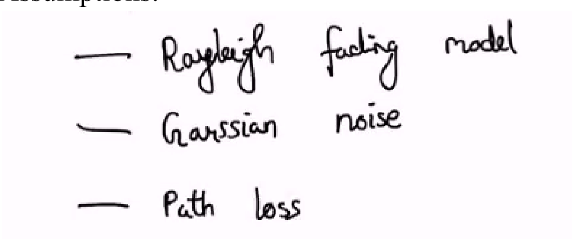
The most difficult way of doing this is called a real experiment when we move around with a smartphone or laptop and do the actual playing of the video. This is expensive because you probably need to hire or rent a car and also spend a lot of time because you need to actually need a vehicle to collect and run the experiment.
Instead of you moving around, can we check the bandwidth traces that your phone is receiving? If you collect that as a file then you find some technique of replaying that file. That is called emulation.

That is closer to reality. There can be hidden factors as your results cant be 1–% reproducible. If the latency is not governed by the last file of bandwidth, then bandwidth trace gives misleading results.
We want to spend less time and want good results, we go for simulation, emulation, then the chances of mistakes are higher.
Using the software you can artificially add latency and artificially construct Bandwidth.
Understand the bandwidth and latency over a longer period of time. You connect your device to a network with very good connectivity to a wired network(Ethernet). Use some software tool to artificially add latency and reduce to bandwidth to exactly the value which you saw in the traces. You are doing the experiment in a more controlled environment, there is a scope of making a mistake here.
The main scope here is supposed you are assuming that the bandwidth trace/latency trace is responsible for the video calling. But it is possible that when you were moving around, during the user's mobility, the server load was actually reducing the quality of the video call. So it is not due to the bandwidth trace. This cannot be captured using emulation.

If I want to do the same experiment through simulation.

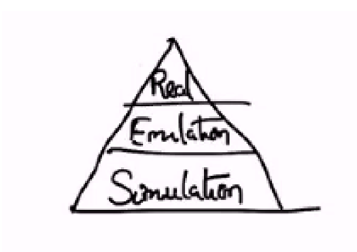
SIMULATION(example)

There comes tools like NS3, your own simulator. This does not collect the real dataset at all, this uses some fading model to make an assumption around the environment. It will use some mobility models which might be stochastic in nature to model the user's movement. By simulation, you are depending on these assumptions. Even when we do some work on simulation on WIFI 6, we assume fading models.
Assumptions:



The path loss does not depend on the atmosphere conditions, but it depends.
Also, the fading model depends on the height of the building but we did not consider that.

They deploy on simulation, and then after actual understanding, they deploy directly.



Real is hard, but most realistic.
Emulation is moderately hard but less realistic.
Simulation is very easy and depend on assumptions and models and is not at all realistic.

When you do real experiments, they are ok with a few experiments.
Supose we have a wireless AP. We wanna evaluate by what will be the amount of latency and throughput if we increase the number of users.
The expectations on the number of users can reduce in the user devices.
If we are doing some wireless sensor networks, in practice there can be 1000 nodes. That will be too costly. In the simulation, you can do 1000 nodes
Real deployment has some value.

Because the DL and PL are implemented in hardware that is why simulation is done is practice. It can be done is software, it can be done using SDR.

Software Defined Radio(SDR) is a transmitter/Receiver but with parameters controlled by software from a computer. Usually, we have these, but this is very expensive. 2L to 10L is the cost of the software-defined radio.
Parameters are the modulation, bit rate, carrier frequency, all of this can be specified. It is providing mode interfacing and more close to the real experiment. Like we have a phone, we can transmit signals using the phone. It is decided by the hardware. Nothing is controlled by the hardware.

Software-defined radio is needed for real experiments.

It is a transmitter/receiver, but with parameters controlled by software from a computer.

→ Rs 200,000 to 10,00,000 cost

Smartphone → Modulation, frequencies are all decided by the built-in chip.

Another issue is the moment we do something in software, we tend to lose performance. SDRs are often slow and have poor performance. It is still a challenge in today's day. It has to go in the USB cable, then etc.
They look like a chip having an antenna and using radio waves. They have to be connected to the laptop using a chip.
It is burnt under it. It is a microcontroller.

MAC layer → Controls access to the medium
There are two types. One is called distributed or decentralized.

MAC Layer → Controls (Arbitrate) access to the wireless medium

Two types → 1 Distributed / Decentralized

WIFI

Another one is scheduling based or centralized.

2 Scheduling-based / Centralized →

5G, LTE

Other techniques use a centralised mechanism, GPS, SDN controller the centralised mechanism. SDN is not at the MAC layer. It is not used in that sense. There is a concept of wireless SDN.

Is Wifi decentralised? YES.
It is not centralised as of now.
Wifi is like you connect the device machine to the backbone and it starts functioning because of it's decentralised nature. There are a lot of techniques by which this can be managed. It is done using contention-based protocols.
The main idea of the contention-based protocol is that it tries to use some sort of randomness to decide whether a node should transmit or not.
Randomness is required because there is no direct control, the same medium is not used by multiple devices at the same time. The easiest way to transmit is Aloha

Aloha → When we have N nodes/ N users and each node is transmitting some data with some probability P.

Aloha is not used in wireless communication. In wired data link, we can detect whether the collision is happened. That is why you know someone is transmitting, then you can back off. In wireless communication, we do not know the collision has happened or not. We may know the acknowledgement has come off.
The exponential backoff algorithm is also there in wireless communication.

If we want to analyse Aloha in practice, assume that each user has a packet to transmit. This assumption is saturated traffic where we use the simulator).
Now suppose that each user has a single packet to transmit and you can also say that the user take the same time to transmit the single packet. Then the probability that 1 user succeeds in transmitting.

$$P(\text{1 User succeeds in Transmission}) = P(\text{N attempts, but only 1 user is transmitting})$$
$$\sim B(N, p)$$

$$P(\text{1 User succeeds in Transmission}) = P(\text{N attempts, but only 1 user is transmitting})$$
$$\sim B(N, p) \qquad \binom{N}{1} p (1-p)^{N-1}$$
$$= Np(1-p)^{N-1}$$

We want to chose p in such a way to maximise B(N,p)
What should be the value of p such that you get the highest number of successful transmissions overall?
Then chose p=1/n

$$= Np(1-p)^{N-1}$$
$$\simeq \left(1-\frac{1}{N}\right)^{N-1} \simeq 0.18$$

For large N, it comes to be 0.18
Aloha can work if N is small.

In the wired medium, we can listen to the medium before we transmit.
Think of ethernet network, we do not have centralised access. We have heard of the token ring network.
This also follows centralised access.
If we go in wired, Ethernet is the most common distributed wired network.
Idea is that you can transmit whenever you have data. But, after transmission, look at the channel's power or voltage levels. If it is higher than a threshold, then we will wait for a random amount of time before transmitting again. Because if it is higher than a threshold, then a collision has happened. That means two of more have transmitted at the same time.

We will then wait for a random amount of time and then we will retransmit.
Because everything was connected over a single wire, then we are able to detect the power of voltage level.
Ethernet still uses this technique in wireless networks, we cannot identify this. Wireless networks, always ask for acknowledgement at the MAC layer itself.

Actually, TCP acknowledges differently. TCP is at the transport layer. MAC layer has separate acknowledgements. They are not related to each other. The TCP acknowledgements are like Payload to the MAC layer.

There is a thing known as packet inspection. It takes the header for consistency. These data about the MAC address are present in the header. If you keep doing this checking, then you will find that whatever data is being sent has an incorrect MAC ID. It is possible to figure this out by consistency check.
All security things will bring some overhead. How to reduce overhead, that is some other challenge. Can these techniques be deployed on the routers? Yes

We are sure that the collision has happened so you can avoid it. The wifi uses a slightly different technique, the power is lower, we cannot check a collision is detected.
Wifi does collision avoidance. Even if there is no collision, wait a random period of time and then transmit. This randomness will reduce your chance of collision even though the other nodes are transmitting.
This is actually a very important idea in terms of Wireless MAC protocols, we should wait even before transmitting, the amount of maximum time we must wait before transmitting is called a Contention window. It can be a number of 5. CW=5
Use a random number between 0 and 5, and you get the value

CW = 5

val = random (0, 5)

wait for val slots of time before transmitting

Check whether an acknowledgement is received or not. If so, then repeat the same process for the next packet. If no acknowledgement is received, assume that collision has happened and double to the contention window.

CW = 5

val = random (0, 5)

wait for val slots of time before transmitting

check whether ACK is received. If so, then repeat the same process.

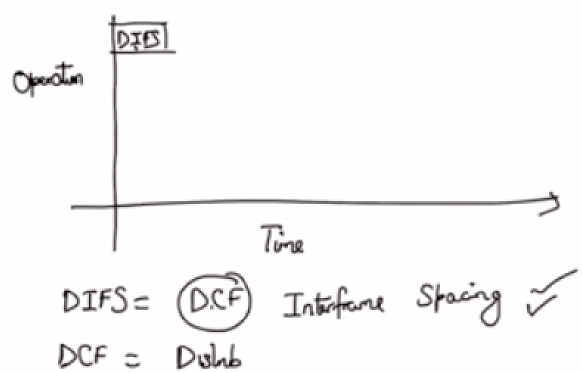
If no ACK is received, assume that collision has happened and double the contention window.

TCP works poorly if the packets are lost. At TCP we deal with 2% packet loss. The moment we observe TCP is seeing packet loss, it is because of congestion, it reduces the amount of packet at a drastic layer.
At Mac Layer, we can handle the 30% losses. This is a poor performance. TCP has the scope of optimisation.
TCP has to handle all those cases. It is still difficult. People still study the performance of TCP under different challenges.
TCP works at the ms level. MAC layer works at sub ms level, usually less than 1ms time level. If you are doing something in hardware, then it can be fast. We should not use TCP.
UDP delay depends on MAC layer delay.

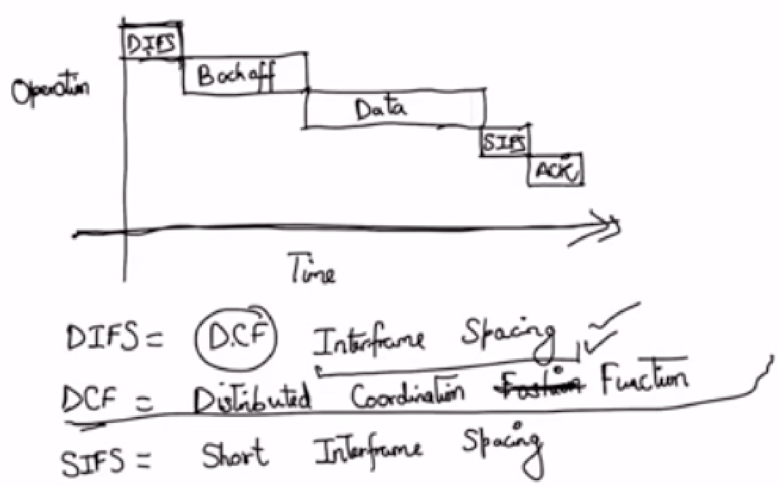
Usually, there is a limit, CWmax. There is a fast restart in TCP.

Lecture 10

MAC layers of Wifi
Let us see how does carrier sense multiple access works in practice
In general, CSMA CA means carrier sense multiple access with collision avoidance. The way collision avoidance happens is through the random backoff. Instead of transmitting immediately as soon as it is available, the user device waits for a backoff period. In the case of Wifi, this turns out in practice is in this form.



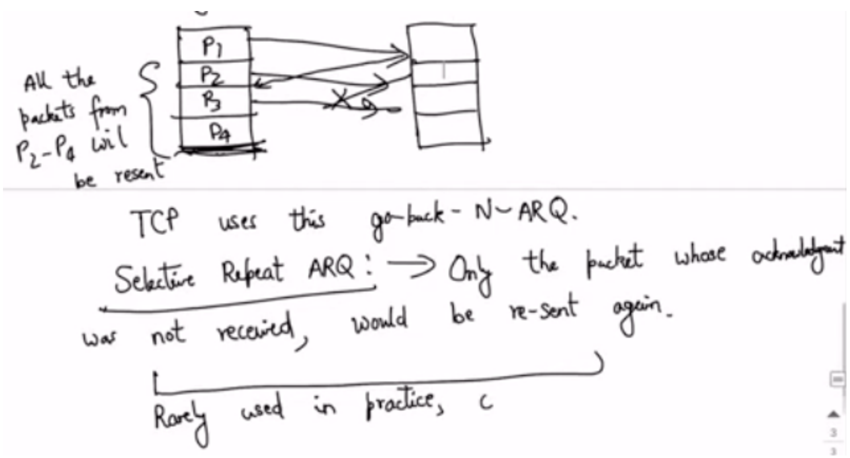
If we go deeper into how WIFI works, then we drew this time axis to understand how a packet is sent, and on the y axis, we have an operation, the first thing that happens is DIFS time, it is DCF inter-frame spacing. Basically, this is the amount of space each successive packet requires. The channel access is being handled in a distributed manner. There is also a way of doing it. Coordination etc is there. Initially, we have the DIFS because we need to wait until another packet has completed transmission. Once you are done, we have this backoff time. We cannot send it directly without waiting for the backoff time. Because there is randomness, in this way you are able to avoid the collision, then you are sending the data. Once you send the data, the data has to be processed by the node, to understand whether the data has arrived in its proper format or it is not in the proper format, might be there are channel errors, the recipient nodes need to understand this. SIFS stands for short interface spacing. . the recipient sends the acknowledgement packet.



DIFS → wait for the previous packet transmission and acknowledgement to be completed.
Distributed Coordination Function Inter Frame Space.
BACKOFF → Wait to avoid a collision.

DATA→ sending the data
SIFS→ Processing the data (short Interframe spacing)
ACK → ACK is received.

What is ACK do not work?
There is no system of negative acknowledgement. If data does not come then its acknowledgement does not come. The sender should infer that the data has not been received. The sender should retransmit. We made an assumption that the sender understands whether the packet is received or not within a very short period of time. The assumption is only valid in WIFI. This is Stop and wait for ARQ. ARQ is an automatic repeat request. Wifi uses stop and wait for ARQ because it is suitable only for short-distance communication. If you need something like long-distance communication over wireless, then in long distances, a lot of time would be taken out before the acknowledgement is either received or not received. ONCE YOU send the data, you have to wait to understand the ACK is received ot not, the **number of DIFS will be very high.** Most of the time will be spent waiting, so this type of ARQ is not suitable. This type of stop and wait protocol is not used. So the alternative is to GO BACK N ARQ. The idea is simple, even if one acknowledgement does not arrive. Then also keep sending a number of packets. Then, check the last packet which was not acknowledged. Resend the packets from that point. It will send all the packers from P2 to P4. It is also done in the upper layer. TCP uses this type of GO BACK N ARQ. There is another system of ARQ, which is Selective Repeat ARQ.
Selective repeat: Only the packet whose ack is not received.



Wifi uses **Stop and Wait for ARQ.**

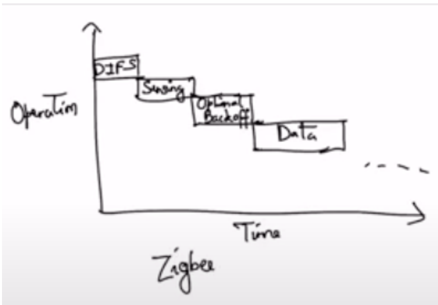
Carrier sensing is not done at the physical layer. Instead, there is a system of virtual sensing, implemented at the MAC layer itself. Whenever someone sends the packet, it should not happen that the user device node, keep sending all the time, then there would be hurt in fairness. The same should not be sent. This is enforced using this network allocation vector(NAV). Whenever the node sends, it starts a counter. Until this counter becomes 0, it will not try to send again. In this way, it can actually modify itself. There is no one node and user which keeps on sending repeatedly. Depending on channel conditions, the value may change. This is the way the counter has to govern the system so that a single node is not sent repeatedly.
Carrier sensing is not possible at the physical layer. Why?
There are two possible reasons. When you are transmitting, we cannot really sense the channel easily that a packet is coming or not. **The transmission and sensing is using the same radio part.** The second reason is that when we are sending wireless packets, unlike a wired channel**, just by adding the power, you might not get sufficient accuracy to understand if anything is being sent or not.**

This is done by **Zigbee**.
The wifi is 802.11. The wifi is adding mac layer latency. Cant we do something like we just send packets at higher power, just to do carrier sensing?
Zigbee is 802.15 and zigbee does that.

The wifi has overhead, but it is cheaper. Zigbee is more expensive. Zigbee is a protocol developed for lower complexity devices. Why is it more expensive?
It is created for IoT devices. Its complexity is less in the software part. Its radio is somewhat more expensive. Because of this ability is embedded everywhere. Many IoT devices also have WIFI. It is very cheap compared to other protocols.

Two types of sensing → Physical sensing and Virtual sensing.

Zigbee architecture

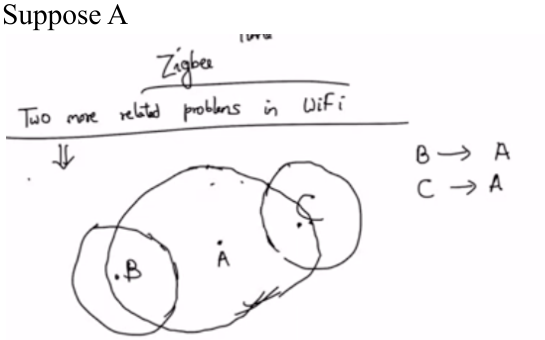


DIFS→ Sensing period→ Optional Backoff → DATA →ACK

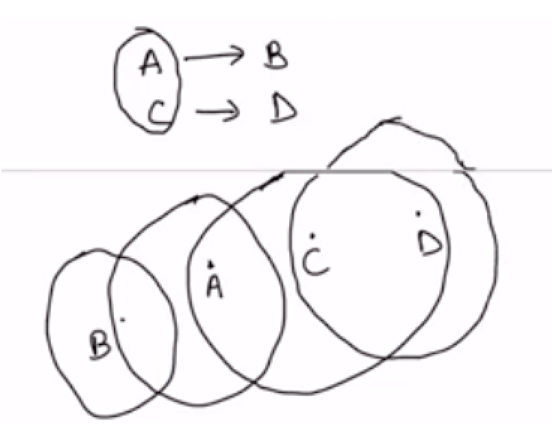
All these acknowledgements are not there to solve one problem basically, there are two more related problems in Wifi, one is this hidden node. This wifi uses broadcast medium, whatever the sender is sending, it could be potentially seen by everyone. We have this issue, that the sender might not know that the receiver is getting data from someone else.

B and C want to send data to A.
B→A
C→A

Effectively there is a very high chance that B and C are not able to figure out that A is free or busy, they can easily transmit simultaneously.
Suppose A



B and C are not within the range, so repeated collisions happen. So, CSMA CA cannot solve this problem. Even if the medium is free, it does not imply that the receiver is free. This is called the hidden node problem. There is another problem, the opposite case can also happen.
The sender is repeatedly finding that the channel is busy, but the receiver is available.
Suppose your sender wants to send to B, and C also wants to send to some other person. A and C are within the range, but B and D are not in range.



A is sending to B and C is sending to D, there should not be an issue, but what is happening is that both of them find that the medium is busy, in practice both B and D are available(can receive packets). There is no need to unnecessarily back off. This is called the Exposed Node problem. There is a good solution to the hidden node problem and there is currently no solution to the exposed node problem. For the hidden node problem, we have a solution that is request to send RTS and clear to send CTS.

The way this RTS/CTS work, instead of waiting for sending the packing, the idea is ,Sender Sends an RTS packet addressed to the receiver. If the receiver is free then the receiver sends a CTS packet. This is a control packet. Since this is a very small packet, so a collision with RTS/CTS packet, will not waste much time. Whenever other receipients see this RTS/CTS packet, they recognise that some transmission is happening in the vicinity. They understand that this transmission is happening. So they wait for a specific time. In this way, this RTS/CTS can handle the hidden node problem. This hidden node problem creates unnecessary collisions. And the exposed node problem does not create any collisions but leads to loss of throughput. So, effectively these are the problems, that happens when this type of activity. Depending on the environment, you need this RTS/CTS. These are widely used in Wifi, and these also have problems. RTS/CTS is required in dance deployments.

What is the tradeoff of using it?
What is the potential loss of using it?
You are sending more control packets, so there is overhead.
Potential gain → Collisions become less expensive since it happened with the control packets.
Ethernet → These problems related to collisions never happened. Why is this ?
Physical sensing is relatively easy, the moment we have a wired system, then everyone is within range of everyone. Effectively, suppose A is very far from C. Then everyone one can sense everything easily. So there is **no chance of any hidden node problem**. The **collisions are very easily detected**. This is all about the distributed coordination function. There are other possibilities also,one of the possibilty is called distributed.
Because the CTS is sent by receiver.
Channel is bidirectional,

```
      RTS            CTS
B  ——------→   A   →   C
               /
      <——---/
       CTS
```

We are optimising the loss. The loss will definitively be there. Collisions will be there, but on control packets, instead of on the data packets.
ARQ techniques are to handle channel errors or collisions. One alternative to handle errors is called error correction coding. Forward error correction is some type of redundancy
Forward error correction → Some type of encoding in which there is redundancy in the data. It uses coding theory, where there is this error detection and error correction codes. Error detection is always used because
At the MAC layer, you need to identify whether the packet is reached or not. But in this error correction, it is optimally used and used mostly in a noisy channel. This is one alternative to repeated retransmission, but it is not used in Wifi or short-range communication. It is used when channels are noisier than in WIFI.

Forward Error Correction → Some type of encoding in which there is redundancy in the data.
→ Coding theory, $\underbrace{\text{error - detection}}_{\Downarrow \text{ Always used.}}$ and $\underbrace{\text{error - Correction codes.}}_{\Downarrow \text{ Optimally used, used mostly in noisy channels.}}$

The wifi has an optional PCF(Point coordination function). Error correction is used at the MAC layer, mostly used by satellite communication. It is not used in practice.
Everything was happening in a pure distributed form, but it need not be that, what can happen is that the node, might decide among themselves, how they want t to transmit by assigning some time slots. This can be in the form of a token ring mechanism. The token ring is the most common way. What this implies, is that you have some nodes like A, B, C and D. And at each point in time, these devices are coordinating among themselves. Each device gets the token in a round-robin fashion. This is the way this works. We get the token in a round-robin fashion. Once A gets the token it gives it to B. Only when it has the token, can it transmit the packet. It is never used in practice because transferring the token can be a problem. In any wired protocol, at the MAC layer, if it is a closed range of each other, the advantage is that collisions can be completely avoided, so the protocol becomes much simpler. PCF is rarely used.

There can be some hybrid mechanism and it is called, HCCA(Hybrid Controlled Channel Access)
Wifi has an access point(AP), so the AP, after receiving all the packets, from all the stations, then coordinate and inform that which station should transmit.
This is an alternative way of resolving hidden node and exposed node problems. This is used in the case of the Wifi Mesh protocol.

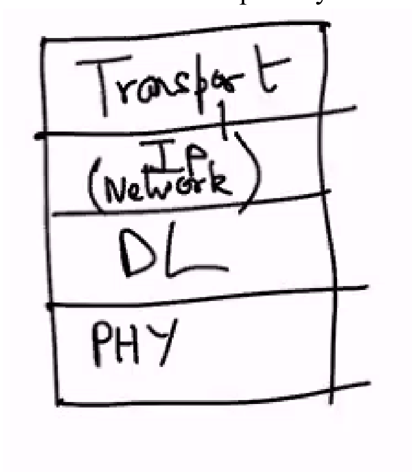
It is very good for throughput, in the hidden node even when the number of users/nodes increase.
Latency of access can increase with the increase in the number of users.
This HCCA is not used in WIFI, because WIFI is evolved with PCF with RTF/CTF mechanism.
Suppose, there is a protocol where the signals are very directional.

This can happen with visible light communication. Laser lights can be directional. Laser has this property that you can have a directional connection. The hidden node problem becomes more severe. In those cases, especially if the channel is not so noisy, then you can use HCCA.

Lecture 11

Assignments and problem questions.
The assignment will require some socket programming.
Go through all the edits of all the classroom posts for references.
Physical classes after the midterm.
15 MCQ in 90 minutes (open books, search online, can not talk to each other during the exam.
A mix of numerical and theory.

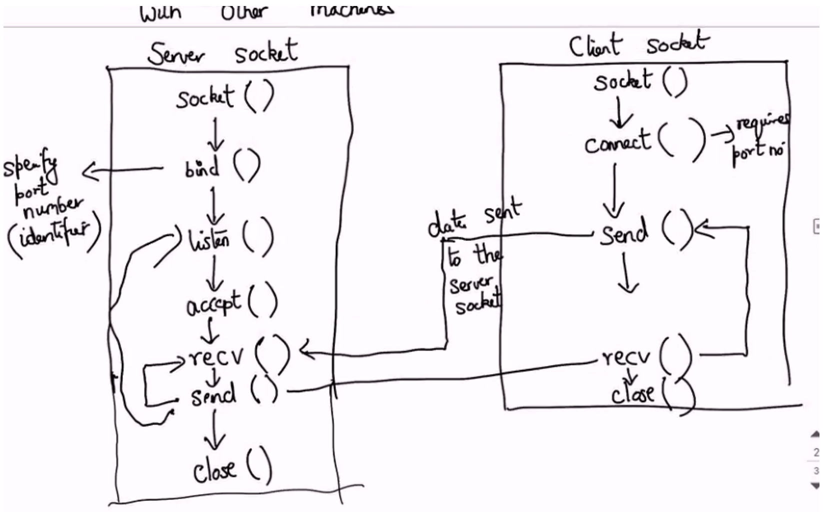
Socket programming:
How to use the transport layer?



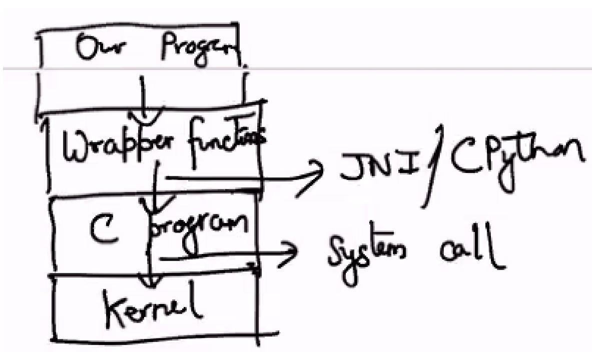
Network and Transport protocol is present in the kernel. To communicate with the kernel, we require system calls. The standard way of using system calls is to define a few functions For reading, you have a read functions. Using sockets, the socket is an interface system called to communicate with other machines. When you use this type of socket, then you have this type of issue that you . There is a two way communication. In the world of sockets, we have client and server sockets.

Socket $\rightarrow$ system call or an interface to communicate with other machines

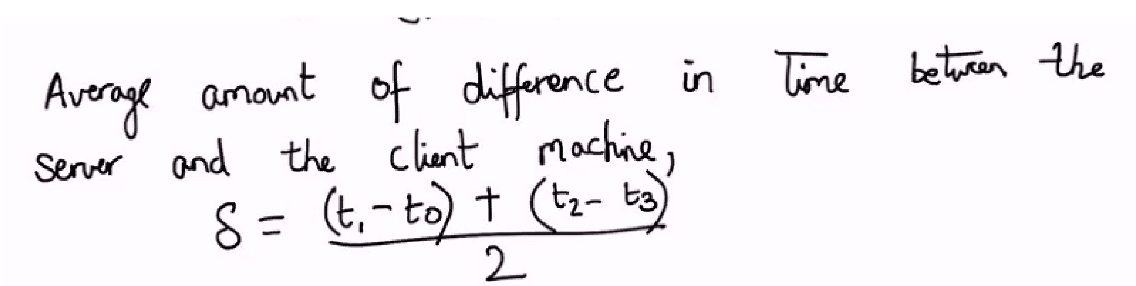
Listens for any request to communicate, and if some request comes, it accepts the request, then it receives the data it receives using the recv function. It can also call send system call.
Using bind, we need the port number at client, we need to specify the port number.
The client socket, has connect and then send sends the data to the server socket.



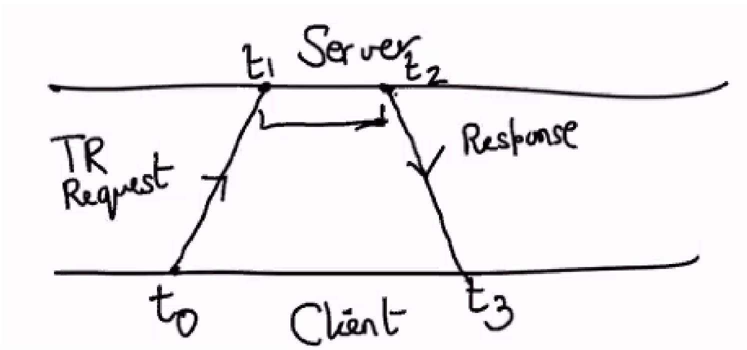
All the functions are system calls. System calls technically in the linux kernel can only be used from C and assembly. Socket library provides a wrapper function so that it can be used in higher level languages. Technically, when we are using Java and python.

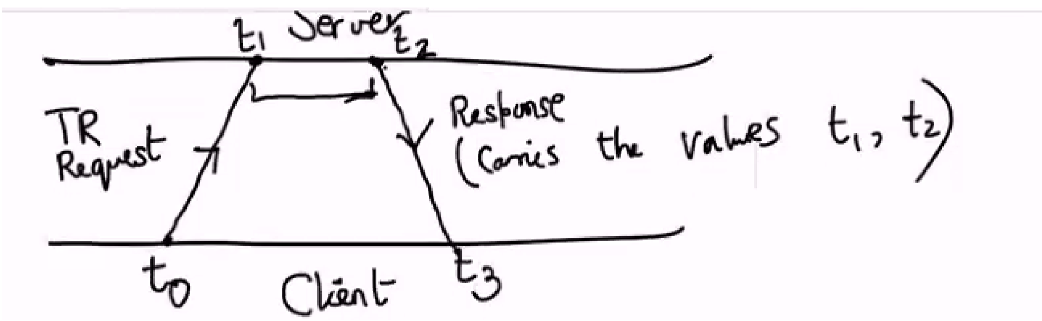


Find out the latency in both wired and wireless. Allocate time on the system. Many of our devices such as phones or IOT devices do not have an accurate clock. PC has more accurate clocks.
Our computers are exposed to different temperatures. The home wall clock may lose accuracy. Having an accurate clock is expensive.
How can you maintain time?
In general, if you want to manage time, then the latency calculation requires time.
Network time protocol (NTP) manages the time. If you send a packet here at t0 and got at t1 and so on. Then the average amount of time difference between client and server is delta.

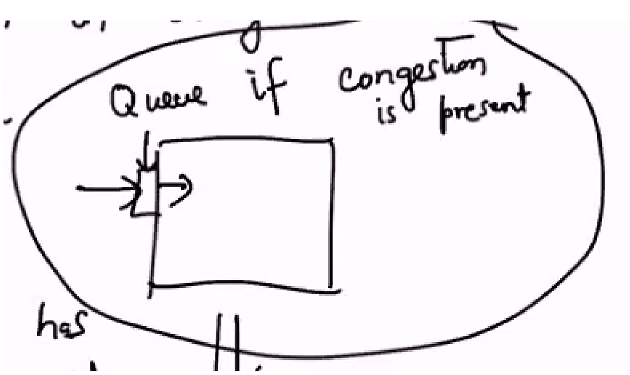


If $\delta = 0$; then $CT = ST$
If $\delta > 0$, $CT < ST$
If $\delta < 0$, then $CT > ST$





PC clock is more expensive. Material is more resistant to temperature changes. The second problem with the phone is, this type of mechanism depends on the channel being symmetric. If this t1-t0 > t2-t3, then NTP will break. The latency is noisy in nature, sometimes the packet takes 10ms. 15ms, 30 ms and so on. Whenever we use wireless, NTP sync will get weaker in wireless networks.
Software noise due to deviations in latency occurs in wired networks.
This time t1 technically should be the actual time when you receive this packet, then we know that I receive this time and I am sending at different times. The wifi card is receiving and then we are sending to OS and then we are losing some time between and so on.
Queue if congestion is present. This queue can hurt your accuracy of the time. The solution is there to this problem which requires hardware support, it is known as precision time protocol which is available on ethernet. The card has the ability to add a timestamp.



GPS itself is a protocol it does not use NTP or so.

GPS is designed in such a way that it does not broadcast the time directly, it tries to explain the satellite locations, and from there on, we can infer the time. GPS synchronisation uses a technique of time difference of arrival. It does not use basic synchronisation techniques. NTP is not used.

How does the server get synchronised?
There is one single atomic clock that this NTP system has and within this NTP, there is a number of layers in NTP. We call it level 0. Level 0 is the atomic clock and at the other levels we have level 1 and then we go to level 2 and usually, we synchronise using level 3 clocks. NTP allows a total of 15 levels. The problem is the more you go down the level, due to NTP accuracy because of symmetry, drift increases. Over wireless networks, drift can be upto 10ms. If you have an accurate clock in the local network and use wired network/connectivity, then we get less than 1 ms of drift. How can you get the accurate clock in practice? This can be done using GPS. Accurate upto the level of ns.

Ping time is RTT time. If your device clock is inaccurate, which is very common on laptops, and NTP is off then RTT is inaccurate.
NTP started in 1995, at that time internet was not expected to become this large, then we thought scope of more hierarchy.
The atomic clock is there.
PTP is also very accurate. We get some microseconds of drift. It is also quite good.

Why do we care? Why time is required?

You can identify using how much time it took for the signal to come in.
How much time means the difference between transmission and receive time. If you calculate that, then you know the radius of this circle and you know the radius of this circle, and you know the speed of light is 299792548m/s

If you are within the 1 km range, then you have only ms of time.

3×10^8 m/s $\longrightarrow$ 1 km

$$\text{Time} = \frac{\cancel{1000}}{3\times \cancel{10^8}\,10^5} = 3.33\ \mu s$$

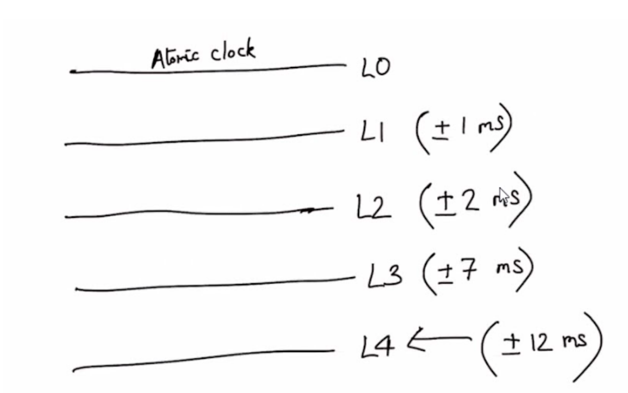
Lecture 13

Discussion of Midterm questions:

Sum of sinusoids, for the simulation, any library can be used.

1. Replace the NIC or Wifi card
2. The same level of the layer can communicate with the same level of layer on the other machine. Data has to go through these lower stacks. The single-packet that won't be received is false.
3. How path loss is related to frequency, we have to understand. Path loss is directly proportional to the square of the wavelength. Attenuation increases with frequency. 24K MHz(5G) signals and this 800MHz is traditionally used by 2G, 3G and 4G. 5G technology would not have worked without beamforming.
4. For the square wave, if we are modeling with the sum of sinusoids. We need an infinite amount of sinusoids.
5. How modulation, SER and BER work. If path loss and shadowing are small, k should increase. Then the SNR will be high, effectively the number of symbols.
6. BER increases rapidly when k increases, due to the limited amount of tolerance. For that SNR, we chose a value of k, we are packing more symbols than are feasible. If you are in a building, attenuation is high.
7. Antenna length should be ordered of km, as frequency is small. The wavelength is larger. Signal will travel slowly. Limited range of the signal. Abc is true.
8. MIMO has multiple technologies, but spacial multiplexing cannot be used without multipath. As they depend on multipath.
9. Use 4 antennae, we did not want CSI to be sent back to the Tx side, we needed some amount of penalty, as we had some specific type of coding, which required sending a pair of bits at time intervals, we are losing on SNR in dB. We are taking 4 consecutive times, we are losing 4 times SNR. It is log(4) it is 0.6. Alamoughty code. Backoff values are chosen using a uniform distribution. Both of the devices chosen it is 1/16 → 4*1/16 i.e. ¼
10. The main problem of rts,cts, is it includes overhead, we are avoiding collision at the cost of sending more control data, we are losing throughput, even if there is no collision, the throughput falls.
11. FDM→ 20 Mhz, OFDM—> 5MHz. 20,28. Whenever we are using guard band, 2Mhz extra, instead of 20 it is 22Hz. If we are actually allocating, there are no guard bands, we need 15Mhz, per carrier wave, if we divide by 15, 440/15→ 29 the answer is not 29 because the last carrier wave cannot be fully accommodated within that. The last one cannot be overlapped with anyone else.
12. Sampling.
13. Scaling in simulation is much easier.
14. Frequency selective fading, an antenna increases, it gets evened out across all the antennas.
15. NTP and PTP are taught in the last class

An atomic clock at L0, and this is the master clock.
When we come to L1, we get a drift of +1 or -1 ms.
Again, when we come here, we add another layer of drift. PTP adds +-5 ms of drift.
When we come to l3, we have PTP, we get 7ms of drift and at L4 we get 12ms of drift. (+5)



NTP supports 15 layers, we can always use more layers.
When drift increases, we have a flatter structure, L11 is sufficient to cover everything. Without transmit beamforming, it is the same as ordinary MIMO. In this cellular, we always have one specific control channel. These layers are not publicly owned, these are privately owned, we need permission to access any of the layers from the owner. Otherwise, we need to synchronise with the atomic clock itself, so whoever provides the OS, provide us with the service.

In practice, they are using the internet. Within LINUX, if you have something internally done, if you dont have internet, set up Linux as level 10 or level 15, it will still work.

LECTURE 14

Details of TCP, including:

- **Flow control**
- **Congestion control**
- **System of ACK and SYN packets**
- **Performance problems over wireless networks**
- **Multipath TCP and QUIC**

Transport Layer

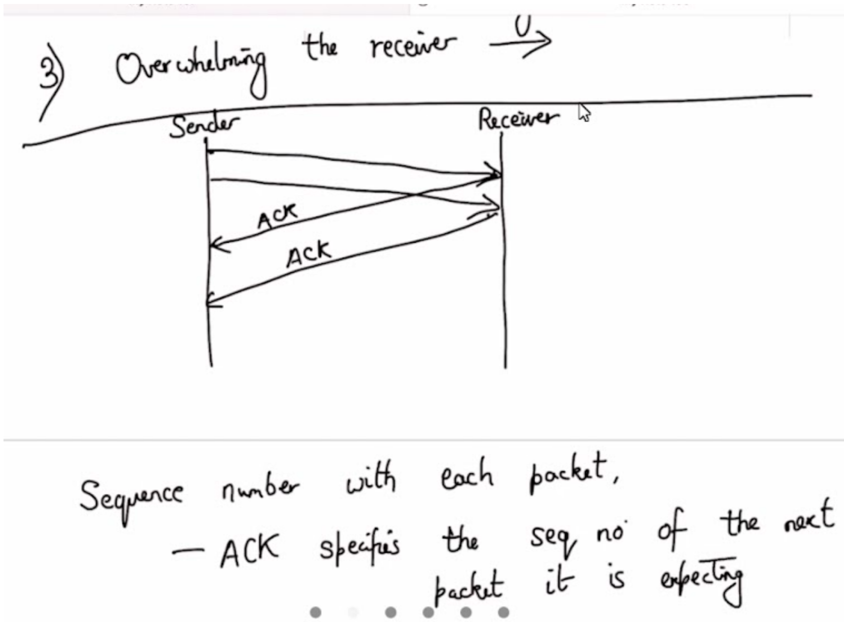
The data link layer is covered, but the network layer is not covered as it is the same for wired as well as wireless networks. The network layer has some weak points, as it does not provide any guarantee of packet delivery. The Data link layer provides depending on the type of service we need, we get some adjustments done for the network layer. Using centralised MAC scheme. We will not see collisions then. Low energy consumption at the data link layer, by switching it off frequently, but it is effect throughput. In practice it is acceptable, we need to focus on optimising the network consumption. The Network layer is just about consumption, the protocol doesn't have any customizability built-in. Internet is running on trust.

UDP and TCP are the most widely used transport protocol. Socket programming using UDP.
UDP provides two basic services. One is multiplexing and demultiplexing which is a very basic thing to do. It provides an address to each process on the application layer. This address in general within the computer is provided in the form of a code number. If you have decided to send this data to port number 1000, that effectively implies that someone has opened. Someone has created a socket and it is in a listening state.

Many standard protocols, their port numbers are already set. 0 to 1000 port numbers are used by existing protocols. SSH is port 22. HTTP is at port 80. Suppose we are using a non-standard service, then port number 1000 to 2^32 for some special non-standard service.
Team viewer does not use any standard protocol. It opens some port number beyond 1000 of its own.
We need to use one port number greater than 1000.
This addressing mechanism → IP + Port Number.

As long we have this IP address, it is given by the network layer, we need which machine we need to go to, withing the machine which socket to go to is given by port number.
The program can use multiple port numbers, because we may have multiple data being communicated at a single instance. We call this multiplexing and demultiplexing, As the computer might be receiving a lot of messages. It is called Demux. Multiplexing is just the opposite, we are sending some data and we are aggregating and giving it to the network layer. The second functionality is error detection using Checksum. UDP is used.

1. Packet losses
2. Congestion →slows up a network for other users, other users are not able to utilise it. We need to use the network in such a way that we are not monopolising the resource by sending too much data.
3. Overwhelming the receiver. → Usually, the receiver has a queue. It stores up the packets.

a. 

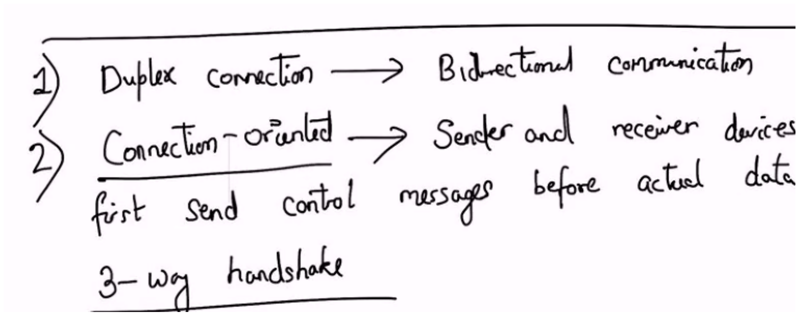
b. You are allowed to send more data, even without understanding the network can take it or not. This can create a problem. Because of this, the standard TCP starts at a slow pace. It is known as TCP slow start. Initially, it will send one packet at a time, and it will wait for ACK. if it reaches the ack it will send twice the number of packets, it will keep on doubling until the threshold reaches. Beyond this threshold, it will not start doubling, it will increase linearly. There is an algorithm, if the number of packets(cwnd) is below the threshold, and you receive the ACK, then you double the congestion window. If your congestion window is greater than the threshold level and you still receive the ACK, then increase the congestion window by linear value. If you do not receive ack, and the congestion window is less than the threshold level. Go back to the congestion window of 1. If you do not receive the ACK, and the Congestion window is greater than the threshold level. Divide the congestion window by 2. TCP does not only have congestion control, it also has flow control, a mechanism to know that the receiver is also sending. There is also a receive window. The number of packets that can be sent between the last transmitted and the last acknowledged. It is equal to the minimum of receive window and the congestion window. min(rcwnd,cwnd). How much you can send without overwhelming the receiver. It uses this type of ACK, if we are able to identify where the packets are going to, the receiver side queue is empty, then the receiver side is empty, we are trying to estimate, how many bytes the receiver can handle, and if the receiver side is failing, from the ACk, he acknowledgement is dropping packets then reduce the receiver window. This is the standard technique that TCP uses. We have four things.
    i. We have this sequence number.
    ii. Acknowledgements
    iii. Congestion window for congestion control
    iv. Receiver window for not letting the receiver get overwhelmed.
    v. Mechanism of retransmission.

c. All this together forms the TCP. There are lot of network conditions where we wanna use it. There is no such optimal TCP. whatever the default version of TCP that I discussed is good. There is a big flaw.

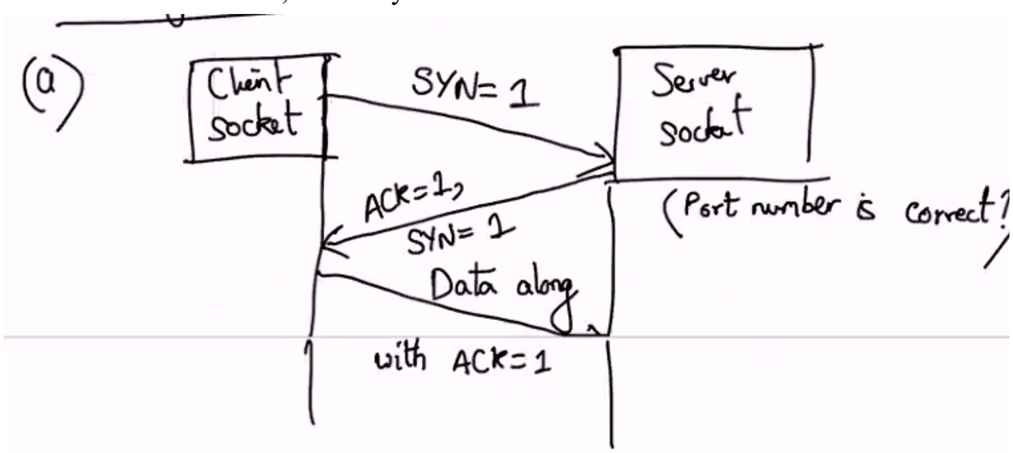
d. Flaw → in the congestion control, as soon as you are getting packet loss, it is due to congestion, in wireless it may be an inherent loss. Like the low signal in the cellular network, TCP assumed it was because of congestion, then we will start reducing the rate of communication. The user find it even worser. We need some modification so that we can distinguish between packet losses becasye of congestion and that with because of poor channel. There are some applications of live video streaming, even we lose packet. What is the situation, it has a lot of redundancy, if one pixel get lost or group of pixel get lost, suppose something like you delay the entire series of packet because of that, that is actually worse adn people will see late, it is better to see inaccurate rather than seeing it late. That is why TCP is not a good protocol

e.

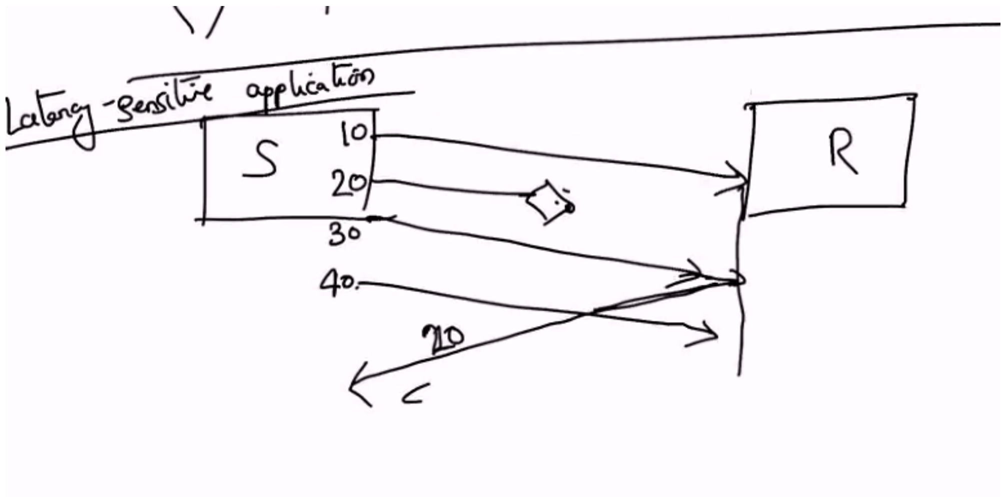
TCP

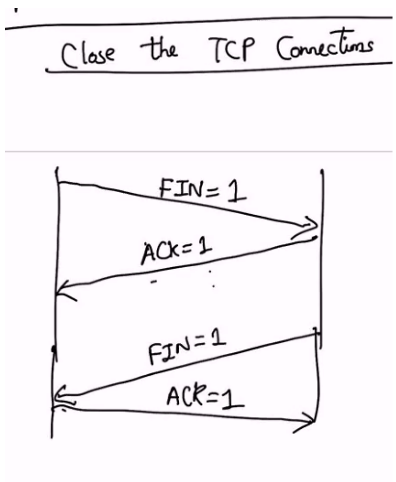
1. It is duplex connection → whenever two devices are communicating, bidirectional communication is possible.
2. It is connection oriented → Sender and receiver devices first send control messages before sending the actual data. It is also known as 3 way handshake.
3. 

4. Client socket initiates the request and it sends a request with a specific flag on. The TCP packets are organised, there is a SYN flag. When SYN=1, the server socket recognises that someone is sending the request. Port number is correct or not, then only server sends the ACK=1 as well as SYN=1



5. Now, both of them know that the communication can take place from both the ends.
6. Suppose one of the devices is not ready, at that point, there is a system of timeout. Suppose one packet does not come in right time, then it will keep waiting and it will go to the first packet again. There is a system of time out. It is known as three way handshake as there is a need of three messages.
7. Advantage of this is→ This type of set up allows congestion control.
8. Disadvantages → latency is high initially. When we are doing web browsing, at that point, the web browsing uses the establishment of TCP connection. When we are trying to establish TCP connection, now a day data is coming from multiple URL, then latency becomes high, we have more focus on page load time. Another disadvantage is performance over wireless networks can be poor. Non latency sensitive applications have use of TCP. If we look at the system of acknowledgement,

9. 

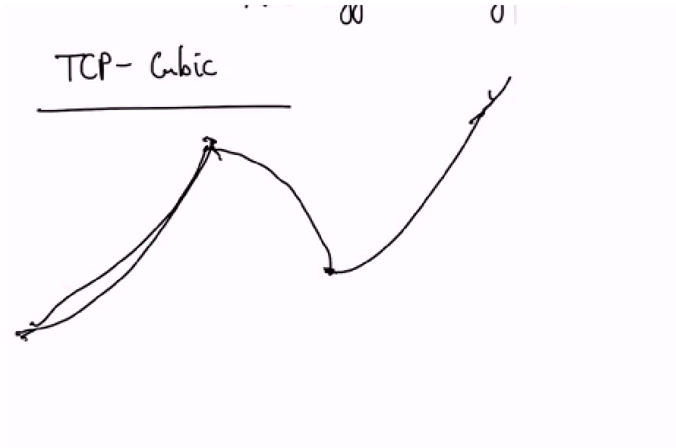
10. Performance over wireless can be poor, in wireless network, loss of packet need not mean it is congestion. There have been some works done, to optimise over wireless network. One of that is called as fast retransmit. Usually, we have a system of timeout. We try to estimate the RTT. We estimate the RTT that how much time is needed for a packet to get acknowledged. We can look at the ack time at the beginning of the connection and them we decide the 4RTT is the timeout period.
11. Within this timeout, if we do not get ack, we assume that the packet is lost. TCP packets are going over internet, it can take a long time for the timeout to be reached. 100ms of RTT is not uncommon, 400ms is a lot of time. There is a specific optimisation. If we received 3 duplicate acknowledgements, we assume, that the congestion is not there. It is just ordinary packet loss and we can retransmit the lost packet. If we are losing packets due to congestion, then there will be no duplicate acks. Duplicate packets means, one packet is lost and remaining packets have been received by the receiver.
12. When the SYN packet comes, it sets up the Receive window and the receive buffer. Similary when the sender receives the ACK, it sets up the congestion window.
13. Close the TCP connections. On the sender side send FIN=1. At the server side, we give ACK=1 for the FYN. again
14. 

15. Unidirectional communication can continue for sometime between these two FYN's.

Issues of how TCP has been modified.

Additive Increase, multiplicative decrease → TCP Reno.
TCP Tahoe → cnwd at 1. It is more aggressive congestion control protocol.

Recently, there have been some changes in the way how TCP works. We are using TCP Cubic in modern computers. The rate TCP was converging at the right window was very slow. The lot of traffic was intermittent. At that point, designers of TCP, they used a cubic version. It goes up in a cube form and falls in a convex shape.



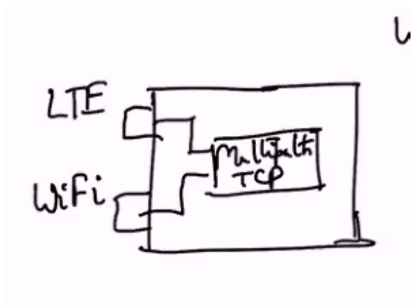
Project wise → TCP cubic.

There are two variations which have come recently, one is actually extention of TCP which is known as multipath TCP.

1. Multipath TCP

When the internet was established, we felt devices will use at most one network interface. But when smartphones came, people connected devices via multiple interfaces. Wifi, LTE, Wifi Direct. Wifi direct does not require any base station. Can we utilise these two network interfaces simultaneously?

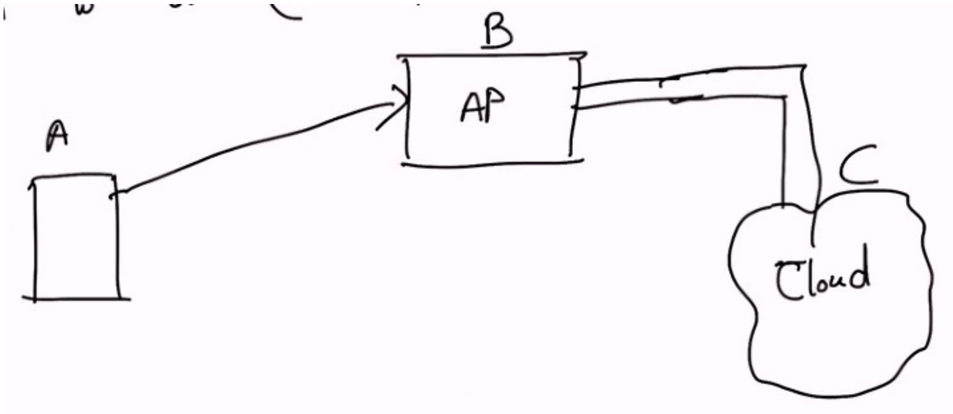
We want to use multipath TCP to draw in packets. We need an intelligent algorithm for how should each packet come in.



Problem: If we request one packet through LTE, but if we found the sequence number which is greater that has reached but it has not reached.
Suppose LTE and Wifi both are being used, suppose LTE latency is greater than that of Wifi.
Suppose we have asked one packet through wifi and it has reached, its previous packet which was requested via LTE, that has not reached. This packet received has to be buffered.
One place it is widely used, it is data centre.
In data centre, everything is in a controlled environment. We can ensure that the network is very predictable and the bandwidth has less variations.
It also depends on energy consumption. Both of the interfaces consume energy.

2. Split TCP

Most people do not use full wireless network and they use wireless network till the base station. And wired network from the base station to internet core.
Centre of the internet is called as the core of the internet.
Suppose we are using our machine to some AP and the AP whether it is a cellular network system or wifi Access point and after that it is connected to wired.



From A→ B we use one TCP connection and it will use optimisations in TCP for wireless networks.
From B→C we use standard TCP.
This split TCP is used in 4G system.

The server tries the split TCP. it buffers and move to different TCP connection and then again sends it.

What can we do more from TCP?
Some applications care more about latency and some cares about reliability and some application are fine with low bandwidth but we need aflow.

Traditionally TCP handled flow and congestion control. Here real time traffic was not there.
Retransmissions and time out came in.

TCP QUIC
Google came with idea that use UDP at th e transport layer but within the application layer, create a sub layer and do the congestion control and flow control at the application layer.
The application layer is split in a sub layer. The lower sub layer can handle the congestion control and flow control. The upper sub layer handle

The transport layer is part of the kernel, it is very hard for the kernel to modify its protocol depending on the application you are using. Now we have part of the application layer, we can use different protocols.
Our network was so used to the TCP packets, the firewall tended to block the UDP packets.

Google meet/ Google applications over chrome.
QUIC has become a separate standard, QUIC

Other issues in the transport layer

1. How does TCP perform when the latency is high?
    a. We have been talking about the bandwidth, suppose we are far away, and the latency is very high. Does TCP work really well? It has been shown that even with high latency, TCP can work well, congestion window can adjust itself so that even at high latency it provides reliable service provided that the traffic is not real time.
    b. Many of you say TCP is reliable. Reliable in DLL is different from reliable in TL. If the DLL and NL are not able to send the packets, then TCP cannot guarantee something. If the data does not actually go through, the sender comes to know it has been lost, it is aware of what is happening.
    c. In most cases, TCP tahoe or TCP reno
    d. Does not work with variable bandwidth.
2. Energy consumption

a. It is not energy efficient. The main reason is because it requires constant sending of packets and ACKS, the lots of control messages are being sent. One of the aspect of mobile protocols. It is a cross layer issue, whenever you send a packet, the wireless interface stays on for sometime.
b. Suppose you send a packet from the sender to the receiver, what happens, instad of the wireless network switching itself, it expects more packets to be send and stys on for some time. The data is relatively small, it will never allow network interfaces to be quick. TCP can be very energy hungry.
c. The clubbing of packets, if whatsapp messages dont get received.
d. When network people talk about energy consumption, it is both on the devices as well as the server.
e. Cell tower must be connected.

Lecture 15

Energy efficiency

Why do we care about energy efficiency?
There is a case for allowing the batteries to last for as long as possible and this is becoming even more true because IOT devices have a battery constraint. We will deal with Smartphone energy consumption.
The key question is before even getting into wireless networks, what actually consumes the energy of the smartphones?

There are processors, Graphical processor, Wireless network interfaces, WiFi, 4G, 5G.
Display component and there are a few additional interfaces like GPS.
Most energy consuming component would be display, but the graphical processors consume the most of the energy , after that display comes, after that GPS radio comes, then 5G, then 4G and then 2G, then NFCm then processors etc.

Why So?

Why cellular networks consume more energy than Wifi?
Cellular networks have more range and energy consumption increases with range because wifi signals are sent with lower range. If there is time, then we will touch upon the display. There is a trait about the communication protocol that the Wifi energy consumption is the least and if you go up the generation, then it is low. This is true, iff you look at energy per unit time for constant time. Isnt this an inappropriate thing to suggest?

Is there a power consumption difference between Wifi 5 and Wifi6
Wifi6 allows beam forming and allows MIMO and that comes with a power cost and that is not a fair comparison because we should compare the performance and energy consumption per bit and not per time. The reason is that it makes a big difference depending on during some bulk download, this 5G will download in less time and 4G and 2G will download in more time. Because it will download for me time.
If you are doing web browning, amount of data is not very high.
We call it as unsaturated traffic. Intermittent traffic is also there, there can be sudden requests from the user.
Then comes this situation. How can you optimise for these cases, because in these cases 5G is not that efficient, Wifi is not that efficient.

Wifi is very energy intensive protocol traditionally, lot of optimisations are there so that it does not consume much energy, PSM came to Wifi, which is known as power saving mode. Wifi chip on your phone goes to sleep after a few seconds of completion of transmission or reception of packets.

Energy efficiency of wireless networks → Smartphones energy efficiency,

If we profile individual components of the smartphones, the graphical processor consumes most amount of the energy followed by the display.

GPS consumes lot of energy. In general the more range, you need high power, energy consumption will increase with higher range. Since wifi has a very short range, then its power consumption is very low.
5G interface sends data much faster than 4G.

Let us focus on wifi part

Why does wifi consumes so less energy, wifi was not energy efficient traditionally, due to lot of optimisations due to which wifi is very energy efficient.

Power saving mode is superb in wifi.
During the sleep, wifi does not work, but the AP keep the packets in the queue, gradually, after a fixed duration, the AP sends a becon signal with a specific bit for each client, if that bit is 1, then the small part of the circuit on the Wifi can recognise it and wake up the entire chip. Then the packet can be sent by the AP. This is what we call as the PSM. This is used in most wifi clients. It is in all wifi clients

The disadvantage of this approach: Latency is very high, packets are waiting on the AP. If your duration is not proper, then TCP retransmissions can happen if the durations become larger. This is the most basic thing that the Wifi does.

1. Adaptive PSM

Your MAC layer can learn by looking at the Traffic pattern what the time interval of the sending becon should be. Depending on the latency sensitivity of the application, you can adapt the time interval. Suppose we find that some p amount of the latency is acceptable then you allow the sleep duration to be some function of p. We can show by some mathematical calculations in the paper proposed. Not everyone have same latency sensitivity. In video conferencing, the sleeping duration should be smaller and in the web browsing, the sleeping duration should be longer.

2. Latency vs fairness tradeoff

Gradually the unfairness comes in, when the Wifi chip using PSM mode, some packets might come in, all the PSM packets should come, if you send the older packets, the latency-sensitive might create an issue. On the other hand, if we chose the new packets, when the client has woken up, the new packets are more urgent. There is a latency vs fairness tradeoff.

When the PSM packets are send first, then the latency sensitive requests might suffer and when the CAM packets are sent first then the fairness of the requests could suffer
Because the PSM packets have come first but are sent later.

3. NAPMan (network assistant power management)

Do not maintain just one queue and one beacon for PSM.
Have multiple queues and different beacons for each client because there is no interference between clients.
Hardware becomes very complex because all the MAC layer logic has to be embedded into the chip. Paper through experiments, we learned that instead of doing it for all clients, if we follow the steps and use 4 queues and 4 different types of beacons, then we get most of the benefit. Gradually, not only this system, if we have multiple access points, then the problem starts. NAPMAN does not work well with multiple wifi access points, because the individual beacons start colliding with each other. Fix patterns when the AP wakes up.

4. SleepWell

It tries to understand and learn the patterns of the other beacons and schedule its own beacons to avoid it. Gradually, it comes to a schedule where the number of collisions fall.

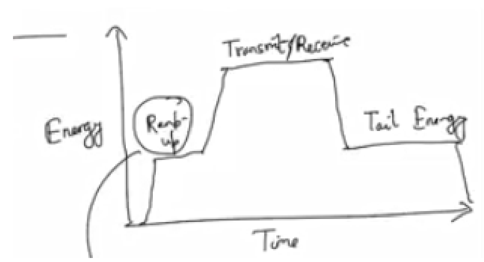
5. Microsleep

It introduces different modes of sleeping in Wifi. If you are in deep sleep, then you can be woken up by using a beacon. You can also have a shallow sleep where packet reception is possible locally but the processing of the packet circuit is switched off. It saves less power but it also incurs less latency.

These are the optimisations due to wifi being a relatively energy efficient protocol. If we compare the IOT protocols that are being used like ZigBee, ZWave and Bluetooth, these protocols consume much lower energy that that of a wifi. The main reason is that with these packets, there is a fundamental tradeoff between latency and energy efficiency and these protocols have higher latency and lower range.
Relay receive the packets and retransmit it, they serve as an amplifier. There we can use all techniques of these sleeping techniques, at each stage of relayed, it can incur delay.
IOT devices like smart bulb and smart lock, all these things, we do not care if it is delayed by 2 to 3 seconds. We can wait for that much amount of time to ensure that it works.

Cellular Network

Amount of time to consume energy has three phases, rampup, transmit/receive and Tail.



We need to set up the channel with the cell tower. Ramp us is used for setting up the channel via negotiation with the cell tower. Once that is done, we will actually transmit or receive the data. There is an overhead ramp up, instead of straight away sleeping, they sit for a bit to see whether a new transmit request comes or not. This is known as tail energy.
TCP has its ACK, the ACK packets can actually come after the tail energy period is gone, The tail needs to be optimized depending on the network. If you know that it is taking longer, then just send a null packet to this network interface to increase the tail time. To avoid repeated ramp up. The other strategy is whenever we have less latency sensitive packets, then accumulate them into the bucket, then send the entire bucket of packets together.
If we are getting messages, we will get the messages at both the application at the same time, my network interface is actually waking up at the same time and bringing them all together,

Data Offloading

Whenever you have access to the Wifi, try to download most of the background traffic. We will see a lot of applications like Google maps.

Computation Offloading

Instead of doing the computation locally, which part of the application os consuming energy. Send that part of the application to the cloud, then bring back the results.
Tradeoff between energy of network and energy of concepts.
If you offload more, then more network energy gets consumed. If you offload less, then more processing energy gets consumed. We should find a right balance.
Privacy Concerns: We might not be comfortable with sending the data onto the cloud.

More specialized hardware:-
Happening very fast on the phone, much of the vector computation we are doing is neural networks, if we have a hardware component which is optimized for neural network. If we have a hardware component only for running neural networks, that can be more energy efficient. It comes from COA. When we are using standard processor, it cannot use

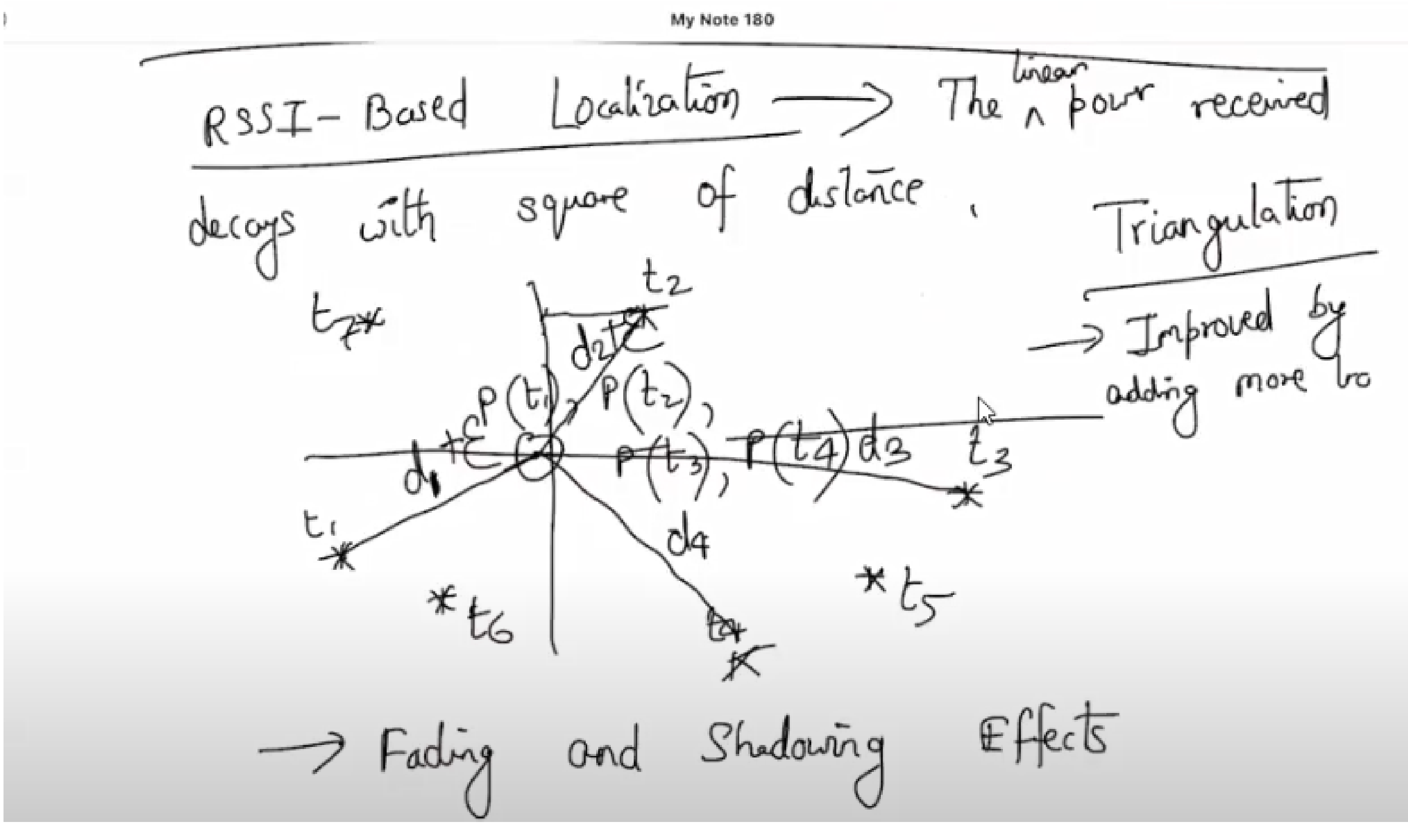
data parallels; within the GPU, we have adders, multipliers etc, the way the GPU is organized, it has to be completely activated for you to run a NN. there is no scope to disable a small part of the GPU and do. If we go to a NN, it is more super specialized, only the part which is required is activated, that is why it is energy efficient.

YOLO → You only look once is an algorithm to recognise the objects. If YOLO is used so often that there might be a case for a specific hardware dedicated to run YoLO. We have to look at the cost to benefit ratio.

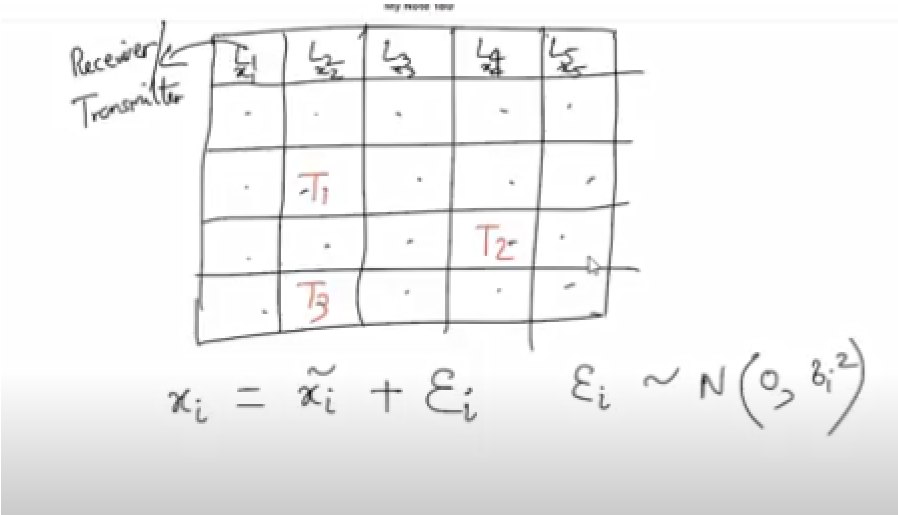
→ Neural processing units on a number of newer smartphones.
Apart from this, there is a display component that consumes a lot of energy. Display comes in two forms on smartphones, one is LCD and the other is OLED. OLED is most common, it has some interesting features. Switching of pixels on the display is identical to setting it black. Effectively if you set the pixel black means not showing the pixel at all.
Many times we have Always on Display.
You want to only check a small amount of information like the time and so on. In those cases, the display uses a few pixels. You can disable colors on the phone. It is known as ultra power saving mode. Disabling colors reduces the energy consumption and you can do more things. Lots of applications that you use. If you have a proximity sensor, then you have this adaptive screen resolution so that resolution is reduced if users cannot perceive the difference.

Lecture 18

Fading and shadowing effects:



Triangulation can be improved by adding more transmitters using a noise term and then using regression. Obviously lots of regression techniques can be used like linear regression, it does not perform well as linear regression may not perform well/ People try gaussian progress regression. This is the basic of RSSI based localisation. Overall whatever it is done, it is acceptable, the terrain is free from obstructions in a reasonable way. If lots of building are there, it will definitely fail, as it does not know where will fading and shadowing happens. There is a sub variety of RSSI based localisation. People are trying more complicate dthing, one basic sub variety is called as fingerprint based localisation where instead you convert it into a more involved ML model. Suppose we have a rough estimate that the user is in this area, we have to calibrate the system,, we have to divide the area into the grid cells. We label each of the grid cells as cell locations. We keep one transmitter/receiver. From each location, wherever the transmitter is, now, at each location, identify the RSSI we are getting. Because RSSI and power values are noisy in nature, we wont get a single value. This X1 or any Xi is summation of whatever the RSSI should be along with the noise and this noise can be different. If we assume that the noise follows a AWGN with 0 mean, for each location



Now, because you have this value, ased on the previous reading, what is the probability of this value coming from a fixed location?

$$x_i = \tilde{x}_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \delta_i^2) \quad (\forall i = 1, \ldots, m)$$

$$\cancel{y_i} \quad P(H_j / y_i) \propto P(y_i / H_j)\, P(H_j)$$

[Bayes rule]

We can remove the denominator, as these are independent events. This is simple bayes rule.

$$L_k = \underset{j=1}{\overset{m}{\arg\max}}\ P(H_j / y_i)$$

This is called as taking the maximum. It is called as maximum aPosteriory estimate(MAP)
This provides the lowest possible error in estimation of it.
Now, we have divided into grid cells and why is it not true that we have not solved it completely.
This localisation problem should be solved completely. Collecting data from each of the locations is very hard. If there is an area in the private premises, and we cannot get into there and collect the data just like that, even though if you collect the data, you can do the localisation locally. We form this grid and in each of them we wrote x1 to x5 and we filled it up and we collected all this data with all the fingerprints we have the idea where the fading is weak and where the fading is strong.

Time Difference of Arrival

This is used in GPS.
GPS is a set of satellites. There are currently 29 satellites and these satellites give out their own location and time. Because we are getting the time at the GPS location nd we know the current time and we are able to find out the time when we wil get the signal, and we can get the time difference across all satellites and using some system of equations.Suppose we find the pseudo random range for each satellite. So whatever we get is

$$\rho_1^2 = (x_u - x_1)^2 + (y_u - y_1)^2 + (z_u - z_1)^2 + \Delta d^2$$

where (x_u, y_u, z_u) is the location of the user, and (x_1, y_1, z_1) is the location of the satellite.

$$\rho_2^2 = (x_u - x_2)^2 + (y_u - y_2)^2 + (z_u - z_2)^2 + \Delta d^2$$

$$\rho_3^2 = (x_u - x_3)^2 + (y_u - y_3)^2 + (z_u - z_3)^2 + \Delta d^2$$

$$\rho_4^2 = (x_u - x_4)^2 + (y_u - y_4)^2 + (z_u - z_4)^2 + \Delta d^2$$

We have a total of 4 unknowns so we need to get data from at least 4 satellites in ordert to solve this.
They are giving you the distance from the satellite to the user.
There may be some deviation if the user is moving,
The estimation can happen in various ways.
Suppose the rough location can be given using cell towers, the these deltas are very well knowon to us.
Solving this is a problem and since these are not linear and these will give the users location.
Solving this is hard as it is not a linear equation.

We can solve this using NR method and we can start by trial and error and then check that in which direction the error is minimised. Newton Ralphson method will take long time to converge.
There is an easier way, we can take an estimate, then $x_u = (x_u' + \Delta x)^2$

U linear equation.

Solve it Newton - Raphson method (which is nothing but start by Trial-and-error, and check in which direction the error is minimized).

$$x_u = (x_u' + \Delta x)^2$$
$$x_u'^2 + 2\Delta x x_u' + \underbrace{(\Delta x)^2}_{\text{Neglect them}}$$

$\therefore$ Converted into a linear equation and then solved.

The more number of satellites we have, the regression error will reduce. If you have more satellites than instead of solving this, we estimate it. Once we have estimate of previous time unit then we can use the regression to get more and more accuracy,

GPS uses CDMA with orthogonal coding.

Because the satellites are very far, the power received at the ground level is very low.
When we get the GPS signal, the standard receiving antennas do not work with the GPS even though the tuning is amazing because the power is so low. We need active antennas. We need some antenna which is backed up by some power. In Active antenna, along with GPS, we have amplifier within the antenna itself. Without this amplifier we cannot get the GPS signal. With any receiving device, it works around 1542 MHz. this active antenna is different from the Gain? This is similar to the Gain but the gain has to be relatively higher. We need 28db to get the signal. The signal is already so weak that we cannot get the indoor signal

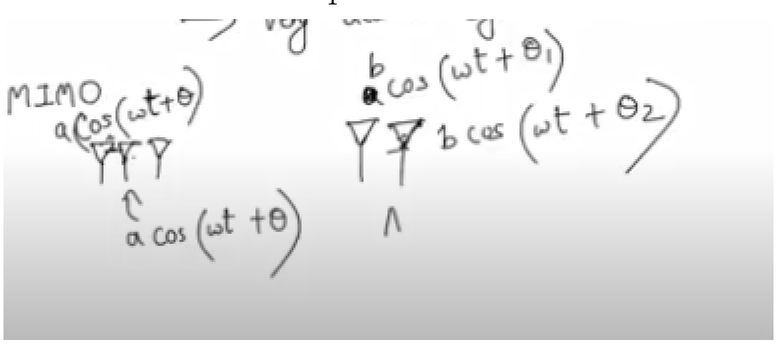
Does the mobile device also have this Active antenna? Yes
Mobile systems actually avoid NR method using cell ID to first to an estimate of a rough location and then find using GPS what is the accurate location. Which cell tower is closest they find the rough estimate of the location. Previously we saw that GPS is also energy intensive.

Can we do Time difference of arrival using cell towers?

The cell towers are relatively closeby and that is why it is more challenging, we need even more accurate system to estimate the time.

In practice, we do not really need to associate the carrier to read the singal, Jio can get the signals of Vodafone. At least this technique is very hard as cell towers are very closeby.

We have a different technique Multiple Transmitters and Multiple Receivers MIMO
Suppose we have array of towers and each of them send specific signal, we see the difference in the angle and we can identify the angle by looking at the signal between the Tx and Rx signal.
For one of them we get Θ_1



Angle of Arrival technique using Phase difference.
TDMA is used in 2G. Here, TDMA had problem that we need more time synchronisation. It is done by 4G, we use periodic sending of pilot signals. They can be used. Effectively using this pilot signals. We can at least use to synchronise the time using this. This technology and synchronisation is still used now a days. This is mostly about localisation.

Indoor Localisation

Wifi Access points

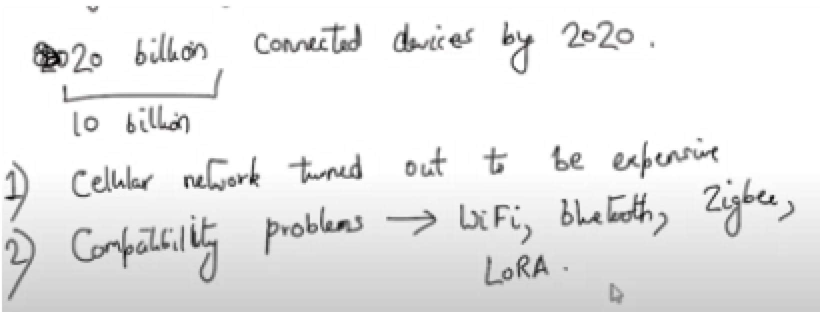
It is same as signature based. Effectively we can use the RSSI values and we can either use regression or calibrate. Question is can we do better?
Currently no technology is mature enough to do better

One problem in Wifi is so many signals coming, localisation becomes harder. People used Ultra wide band localisation where the idea is we utilise some signals of other frequency with much larget bandwidths. Other techniques
We take images of walls adn then identify where the specific signatures are present. To understand this, we can call microsignatures. Overall people are taking two different.
One is based on RF signals, other is computer vision based approach,

There is a computer vision based approach for GPS also.
Overall this localisation has two distinct approaches.

Lecture 19

We will discuss about spectrum crunch in wireless communication.
Scale of the problem
We have talked about IOT. Lot of services are there. When we connect physical things to internet.
Ericsson said 20 billion connected devices will be there.
How many connected devices we have? 12 billions.
India has 100+ smart cities.
Currently lot of smart city technologies is running on wired and not wireless.



3) security problem: devices do not have powerful processors, the technology has to be securely programmed.
DDOS attacks are possible.
Consumers wont install any IOT devices in their home.

Compatibility problem and security problem

Cellular network: More expensive
In india, it is relatively cheap , for any type of communication, we need spectrum, cellular runs on licences spectrum. Idea is by default, the Govt owns all spectrum and the govt. Gives licence to telecom companies based on frequency bands to telecom companies. This is the way it works in every country. These licences are given by drawing bids provided that they are really a smartphone company etc. As more and more consumers are coming in, spectrum getting congested and devices are becoming more and more. They are actually paying, The price of spectrum is increasing a lot. The interesting thing is in a country like India, cost of spectrum is much higher than developed countries. This is a challenge that how to ensure that we are still able to provide good service to the consumer without increasing the price. Question is what can be done, the current system basically does not scale up. The govt has done many thipngs in this space the technology is there. It is about adoption, they have freed up more spectrum. There are frequency bands which were originally reserved for military and police but are not used. This is the easist solution. Second technology has nothing to fo with the gvt. More communication uing higher nd higher frequenc bands. Wifi uses 2.4 GHz. Wifi startedusing 4 GHz and gradually a lot if cimmuniacation. They do not propagatte very well. You need something like mor MIMO. In the US, when 5G was unrooled, then they did not vacate. Even among the newer users, a protocol should specify whether a frequency band is being used or not. This requires a protocol. Effectively there is a discussion.
Imbalance or high difference among utilization of different bands is widely persistent.
The government or any regulatory authority, needs to figure out a way that the end consumers do not have to face any problems.

Businesses need to follow the regulations.
Once you acquire the license, you can use it or not use it. Spectrum will be given based on the idea that only if some frequency band is being used, then other users will not be allowed. If license frequency band is idle and not being used, in that case, other users(users who have not taken any license) will be allowed to use it which means that if you going to use it for something for proving the service, you are allowed, but if you are not using it, then you have to leave it. It requires a lot of techniques and it builds a lot of problems. First of all, already the government has given licence to some of the people. These organizations who have got these licenses are using it according to their own leisure. These incompent users cannot be easily asked to vacate the spectrum. In just the last 6 months, in US when 5G was being unrolled, lots of flights got canceled, then they did not know how to move to different frequency band, they have to cancel the flights.
Leaving off the spectrum, people have to redesign the hardwares and install in all comm systems and ensure theat the range of the systems is sufficiently good. Even among the newer users, there is a protocol, for them to communicate whether the spectrum is being used or not being used. This requires a protocol. Effectively there is this discussion, it is called spectrum access system. (SAS). Whenever any user wants to actually starts to use, they need to inform to the database of what they are sending. There is a database storing the power levels at each latitude and longitude at each frequency. Whenever any user(licensed or licensed) is keen to communicate, they need to update this database. Licenced users do not need to check the SAS database, whether any unlicensed user is using it. Whereas the unlicenced user needs to check whether another user is using it or not.
Licensing implies that yes you can use it without checking if others are there or not. That means you have to update the database that you are using. This is the spectrum access system. Lots of other problems came in. First of all, we need to check what technology is used by the unlicensed users to communicate with?
We can argue that we are talking about long range communication, why LTE is a good approach. Both LTE and WiFi have their plus points. LTE is more suited for long range communication. Using LTE is not very easy. It is about how the voice can be send, suppose the signal quality is not good, you can reduce the low modulation rate always. Long range or short range does not matter here. Currently whatever LTE is there, 800 to 1800 MHz. It has very good propagation speed. 1800 MHz can be a problem in the rural areas, as the cell towers are far.

Second this is how to avoid utilization of the incompent user. The moment we know that we populate the databases, privacy concerns step in. DB is similar to a grid. Effectively if you populate this with the power, P1, P2 and P3. This is the way the SAS database works. You should know when the user is transmitting and where is the user transmitting. This can be a problem for any sensitive communication. This data can be sensitive for any military application.

If you use WiFi, as it is not for long range communication, we can get a hidden node terminal problem, if the range increases. RTS and CTS mechanism cant help us much. Wifi provides much lower throughput over the longer range. LTE has some additional problems, if we are trying to use LTE, then LTE has problems that if LTE is not easy to switch between channels, if we are trying to switch to some other channel, then it is possible and it leads to disconnection for about 30 to 40 seconds. Which is a major disruption to the working of the system. This is one major challenge to use it. People have used a variety of techniques.
One thing that the people have done is they have tried to understand from prior users, which channel is less likely to be interfed with. Main reason why LTE is better than Wifi, is it has centralisied control, but it cannot be completely centralised.
We can talk about semi centralised system, where the channel are allocated to the towers using some algorithmic protocols without direct coordination. Effectively you can say the closest users can get some channels, and everyone knows that who is supposed to get what. There is some mathematical thing involved here.
In the database, who is transmitting, does not have actual idea, how is the matrix filled, there might be obstructions, so no one not all the users who are transmitting do not know how far the signal range they have. They do not know the accurate estimation of the power. How to get the accurate estimation of the power?
This is a major challenge, the easiest thing is to do is to interpolate. Estimating in the remaining locations. Because your primary goal is that you are making a conservative estimate. To protect the licenced user, you are forced to make very conservative estimate. Here we try to deploy spectrum sensors to get better estimate. It means that we will say that the

power here is higher that it actually is. Suppose you have a power at P1, you do not know, you look at P1’. if P1’ > P1 then unlicenced users will avoid using it. That is why it is a conservative estimate. We do not want the licenced user to face any disruptions in the communication. If you do not have a very good estimate, then let me assume we have a conservative estimate. We do not want the communication of the lcienced user to stop. Here the challenge is that we can interpolate and we sort of supplicment the user information by deplioying additional spectrum sensors. Who will deploy these spectrum sensors? Even if you deploy, where should we deploy? Third question is what type of sensors it is. In practice the measurement of whatever spectrum that we can use has type of spectrum we are using. It is 2000 Rs to 200000 Rs to 1000000 Rs.

We use the power of crowd sourcing. When we add a cheap sensor to each smartphone. These smartphones have cheap sensors who have the noisy data. Here is again where we can use this sort of statistical technique to remove the noise by taking in data from multiple sensors. If we had one sensor, then obviously the deploying the extensive sensor is much better, but if we have a series of sensors, then it is more of a optimisation problem. Then deploying more number of cheap sensors, then having more number of cheaper sensors is better then having fewer number of more expensive sensors. Because of the diversity in the terrain, if we deploy the more sensors, even then we have error prone, if we have deployed have fewer number of cheap sensors, the overall error in the first case can be lower. Error in estimate of power in the SAS database is usually lower in more number of cheap sensors assuming we have spent the same amount of money. In one case this is not true, if you need sensitivity at very low power, <-83 dBm> then fewer more expensive sensors is better. There was another issue which I talked about which is the privacy concern. The more I solve the other concerns. Crowdsourcing do not actually make the privacy concern work. More accurate database can violate it. Instead of actually populating the power values. To overcome privacy concerns the SAS database might be populated with some randomness. This randomness - utilisation. There is a tradeoff between how much privacy we can gain w.r.t the utility. This tradeoff can be quantified by mathematically modelling it. This is location based privacy.
This cannot be problem in LTE network.

Spectrum sensor can read or decode the data being sent. The data needs to be encrypted. There might be some regulations. The spectrum sensors should not send some raw data to the cloud severs because this raw data can be utilised to identify it.
Third type of violation → From the type of data the sensor collect, we can identify what device it is. If we have spectrum sensor near you , it can identify that you have changed your phone. Whenever you try to send some signals, if you ask them to be sent at specific power and frequency. Because the hardware is not completely accurate between the power and frequency over which that data is sent. By carefully monitoring, on which specific frequency the data is sent, the sensor can identify that which device is sending the data. This is called signature based detection.

Lecture 20

Spectrum Sharing Technologies
We talked about spectrum sharing technologies which people call as

1. Dynamic Spectrum Access
2. Cognitive Radio networks
3. Let us limit this frequency and some channels were licences using concepts of spectrum shaing. It is called citizen broadband service(CBRS).
4. Tired Spectrum access is allowed.
5. If you use Dynamic spectrum acceess, you need non adjacent channels for communication
6. Usually in practice, tower is not known, If you look at secondary user(unlicenced user)
7. Both of these are challenges. Who will populate the spectrum database?
8. Crowdsourcing, you can attach a spectrum sensor to a phone and then get the details about spectrum usage by collecting the data.
9. More number of cheap noisy sensors give more accurate results than fewer number of expensive sensors.

There has to be some algo which has to be used that can actually combine the results of the sensors and this can be done broadly in two different ways.

One way is suppose you have binary sensors.
Either the sensor senses the channel, then if it finds some power, or that means there is a transmission from somewhere close by then it sends 1 otherwise it sends 0.
This sensor has the probability of detection. In reality, there is a transmission T and my sensor output O is 1.
There is a trasnmittion but the sensor could not detect, this is probability of miss.
There is no transmission but sensor detects the transmission. It is false alarm
There is no transmission but the sensor also does not detect anything.

---

1) Binary sensors → Sensor senses a channel, then if it finds some power / transmission, then it sends 1, otherwise 0.

T

otherwise 0.

	O	T
Probability of detection (P_d) $\longrightarrow$	1	1
" " miss (P_m)	0	1
" " false alarm (P_f) $\longrightarrow$	1	0
" " true negative (P_t) $\longrightarrow$	0	0

If the sensor is proper in nature, then the probability of true negative will be greater then probability of false alarm and the probability of detection will be greater than probability of miss.

If the sensor is proper in nature, then

$$P_t > P_f$$

$$P_d > P_m$$

$$P_f + P_t = 1$$

If the sensor is proper in nature, then

$$P_t > P_f$$

$$P_d > P_m$$

$$S_1, \dots, S_k$$

Sensor "fusion" rule $\Longrightarrow$

Suppose there are k sensors S1,...,Sk.
There is a sensor fusion rule. We are fusing the results of multiple sensors. You need to take a specific value witch is defined by

$S_1, \ldots, S_K$

Sensor "fusion" rule $\Longrightarrow$ Chair & Varshney sensor fusion rule

$$\log \frac{P_1}{P_0} + \log \frac{P_d}{P_f} + \log$$

Suppose each sensor gives a decision D_i

If $\sum_{i=1}^{k} D_i \log \frac{P_d}{P_f} > \sum_{i=1}^{k} \log \frac{P_t}{P_m}(1-D_i)$, then transmission is present

Otherwise, transmission is absent.

Slight modification

If the left side is bigger than the chances are higher that the one that is coming, because it has detected correctly.

If the right dide is having higher value,that means no value has been changed

Otherwise, transmission is absent.

$P_1 \longrightarrow$ probability of transmission being present

$P_0 \longrightarrow$ " " " " absent

$$\log P_1 + \sum_{i=1}^{k} D_i \log \frac{P_d}{P_f} > \log P_0 + \sum_{i=1}^{k} \log \frac{P_t}{P_m}$$

Suppose P1 is probability of transmission being present and P0 is the probability of transmissio being absent. From the axioms of probability, this is the optimal sensor fusion rule.

② Suppose the sensors report the power values, but their reports are noisy in nature.

Assumption: $\longrightarrow$ The noise can be modelled as Gaussian.

1.
2. Adding sensors will increase the accuracy.
3. Suppose the sensors report the power values, but their reports are noisy in nature. Thisi s a widespread assumption, it is obvious.
4. Assumption → The noise can be modelled as Gaussian noise. At any point you have a sensor. That sensor is receiving some value, the power value xk.

Assumption : $\longrightarrow$ The noise can be modelled as zero-mean Gaussian .

$$S_k \rightarrow x_k + N(0, \sigma_k^2)$$

5.
6. Suppose you have whole series of sensors with Gaussian noise.
7. We add all the sensor outputs.
8. We add this sk. Sk will be either 0 or some constant value. I will be 1 to k. Test with some specific threshold tau.
9. If it is greater, then the signal is present, if it is equal then may or may not be present, and if it is less than tau, then it is not present.

Assumption : $\longrightarrow$ The noise can be modelled as zero-mean Gaussian .

$$S_k \rightarrow x_k + N(0, \sigma_k^2)$$

Whole set of sensors with Gaussian noise .

If $\sum_{i=1}^{k} s_i > \tau$, then signal is present

$= \tau$, then may or may not be present

$< \tau$, then it is not present.

$$\tau = \frac{\sum_{i=1}^{k} y_i}{\quad}$$

10.

present

$< \tau$, then it is not present.

$$\tau = \frac{\sum_{i=1}^{k} y_i}{k} \qquad \sum_{i=1}^{k} s_i \sim N\left(\sum_{i=1}^{m} x_i ,\right)$$

11.
12. Sensors errors are not correlated.
13. The probability of misclasification between these distributions is lower than using the single sensor distributions comes from the property of gaussian distributions.
14. You should take all the sensor outputs that are present.
15. Suppose you can only include budget B amount of sensors.

What algorithm can be used to select sensors?

What can be an appropriate algorithm for this?

$i=1$

$= \tau$, then may or may not be present

$< \tau$, then it is not present.

$$\tau = \frac{\sum_{i=}^{k} y_1\, i}{k} \qquad \sum_{i=1}^{k} S_i \sim N\left(\sum_{i=1}^{m} x_i, \sum_{i=1}^{k} \sigma_k^2\right)$$

[Assumption: $\longrightarrow$ The sensor's errors are not correlated.]

$$N\left(0, \sum_{i=1}^{k} \sigma_i^2\right)$$

3

There is a greedy algo, you have k sensors and you have B budget and some function f(T) gives you the accuracy.
While |T| <=B

It is kind of a feature selection strategy.

while $|T| \leq B$

for $i = 1$ to s_k,
if $s_k \notin T$
check $f(T \cup \{s_k\})$
if $(f(T \cup \{s_k\}))$ is highest
then add s_k to T.

return T

This algorithm gives an approximation ratio of $\left(1 - \frac{1}{e}\right)$ times the optimal. $\approx 63\%$

Provided the following providing that adding the sensors only improves the accuracy.

Formally,

(2) Submodularity : →

$$f(T_2 \cup \{s_k\}) - f(T_1) \ge f(T_2 \cup \{s_k\}) - f(T_2)$$

$$\forall\, T_1 \subseteq T_2,\ \forall\, s_k$$

T1 is smaller set than that of T2.

Knapsack problem

$$\text{Maximize} \quad \sum_{x_i=1}^{m} p_i x_i$$

$$\text{subject to:} \quad \sum_{i=1}^{m} c_i x_i \le B$$

Knapsack problem

$$\text{Maximize} \quad \sum_{x_i=1}^{m} p_i x_i$$

$$\text{subject to:} \quad \sum_{i=1}^{m} c_i x_i \le B$$

NP-Hard

FPTAS → Fully Polynomial Task Approximation Scheme

$$\frac{1}{2}\left(1 - \frac{1}{e}\right) \approx 31.5\%.$$

By taking combination of three times, so that you compute O(k^3) comparisons you can get 1-1/e approximation ratio.

Byzantine model →

Suppose your sensor is giving gaussian error, if the noise model is not gaussian, it is very strange, without any gaussian distribution, then the property changes, previously it is always better to include sensors, here it is not true.

Check each sensor, weather its result agree with rest of sensors, if not remove it. It is better to exclude a few sensors who are giving the highest amount of errors.

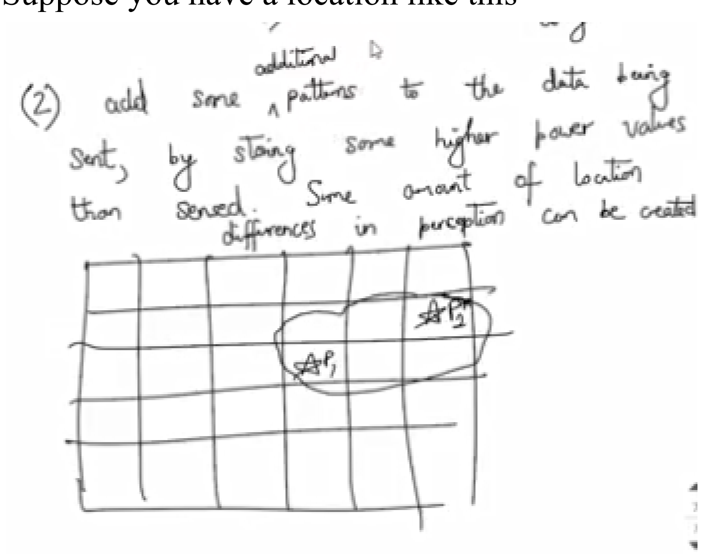
These byzantine model → When we are crowdsourcing, someone who wants to play around with the system, might introduce malacius sensors. It is very important model for crowdsourcing, we need to identify who is trying to disrupt our system. This is one strategy and there can be actually more based on these strategy, you can build new strategy. I assume that sensors are selected in advance. We can use active selection of sensors, afer looking output of one sensor, we select further numbers of sensors, tehre are also some conditions where greedy works well and where greedy fails.

Active selection of sensors is also possible , with both greedy and non - greedy strategies. This is all about mostly impact of how you can do crowdsourcing. This is a wireless network class, on point of view of dynamic spectrum. Crowdsourcing can be used for improving the behaviour of devices, making them context sensitive, fingerprinting can also be done using crowdsourcing.

Crowdsourcing problems

- ➢ Privacy → Can be of multiple types, location based privacy violations can be there, there can also be sort of pattern based. Using crowdsourcing, close to someone's rom during night or daym it is possible to infer user;s network activity. Even we can infer from the sensors that which device users are using.
- ➢ Currently CBRS does not storecrowdsourcing sensor data. There are some strategies proposed to resolve this problem. Occasionally we can add fake reports of spectrum usage.
- ➢ Whenever we sense spectrum is utuilised, send the data to the DB, DB admin knows it is being fetched, You can program the sensor so that it occasionally reports a fke usage.
- ➢ Database dows not know when the secondary user is using.
- ➢ Add some additional patterns to the data being sent, by storing siome higher power values then served.

Suppose you have a location like this



Lecture 21

The moment we get in the perceptions of users, then things becomes complicated, then this makes optimising the quality of experience complicated

Let us understand how to measure perception,

In the case of video, we can actually show the videos to the group of users, ask users to rate the quality of video and then report the average.
Usually there is a standardised,
From 0 to 5 the users will rate the quality of the video.
0 indicating very poor, 5 indicating excellent quality
Because, this is about users opinion, then this average metric is called as mean opinion score.

This has a number of problems
Users participation is expensive.

Mean opinion score increases if the video resolution is increased.

MOS reduces when the rebuffing increases.
Reduces if the startup time of the video increases
Because the bit rate and video resolution can change in the middle of the video, so effectively any change in the bit rate reduces, if the video resolution or bit rate changes in the middle of the video.

How to optimise this whatever may be the network condition
This can be modelled.
Suppose this video resolution → RI
Rebuffering duration →Ti
Startup time → Si

Overall QOE you can create a weighted sum.
Overall QoE = alphaRi + betaTi + gammaSi + delta $|R_i^t - R_i^{t-1}|$

$$\text{Overall QoE } (Q) = \sum_{t=1}^{T} \underbrace{\alpha R_i^t + \beta T_i^t + \gamma S_i^t + \delta \left| R_i^t - R_i^{t-1} \right|}_{\text{Weighted sum of each individual parameter}}$$

$$\text{Maximize } Q \text{ subject to:- } D^t \leq B \quad \forall\, t = 1, \ldots, T$$

$$R_i \rightarrow R_i(D^t)$$

In wireless networks, your bandwidth can keep changing.

Effectively you can write it like summation of all these
Maximise Q

Ri is the function of the data

$$\text{Maximize } Q \text{ subject to:- } D^t \leq B \quad \forall\, t = 1, \ldots, T$$

$$R_i \rightarrow R_i(D^t)$$

In wireless networks, your bandwidth can keep changing.

If the bandwidth changes, then
Suppose the bandwidth fell, you have to encounter this rebuffering.
Suppose the bandwidth falls, then it allows rebuffering.
The resolution also falls down
In general then if we do not have any prediction system. And the bandwidth increases, the system will try to stay

Current video streaming system, do not increase, they keep it same
The netflix is not staying at the same level, they are trying to predict the complex bandwidth.

ARIMA model → Auto regressive integrated moving average model.

Trying to predict the BO by running complex ML techniques.

→ ARIMA model →

$$k_1 B^t + k_2 B^{t-1} + k_3 B^{t-2} + \dots$$

These types of predictions are possible and used in practice in the video delivery system.
In this system, the bitrate is changing, according to the system, so it is called adaptive bitrate system

ABR systems

Can we do better then the ABR systems.
Can we put little bit further
Can we do more than the adaptive bitrate system.

Building any distributed ting is the problem
The video is coming from the cloud server.
Usually the cloud systems they need to keep the video encoded at multiple resolution
Knapsack

Adaptive Bitrate Systems (ABR).

⇒ They need to keep the videos encoded at multiple resolutions and split them up.

Optimisation is temporal in nature.
There is a concept of space related optimisation as well.
Not used in practice → Foveated streaming

Suppose the user is focussing in a particular area.
Then at real time at the server, we encode that at high resolution and encode the rest at lower resolution.

The screens are also getting larger.
It is more easily perceived that what is the bit rate.

Space-related optimization is possible.

⟹ Foveated Streaming → Suppose you ~~foc~~ find that the user is focussed on a particular area of the video. Then, at real-Time encode the video with high bitrate in that particular region and low bitrate in other regions.

There are some challenges here.

How can you find out which part of the video the user should focus on.
Gaze data
Without a lot of calibration the gaze identification is not very activate.
Different users can focus on different

Video content based Flow
Can we actually use the video content which area the user focusses on.
Saliency model
ML model that identifies the regions with the strongest colors.
Users usually tend to focus where the color is the strongest.

What are the problems with the foveated streaming

are the strongest.

What are the problems with foveated streaming?

→ Re-encoding of the video on the server might be required.

are the strongest.

What are the problems with foveated streaming?

→ Re-encoding of the video on the server might be required.

→ Getting gaze data is not always feasible.

→ Used in 360-degree videos

When we are using this 360 videos, they are good to watch if the video is 4k.
We are usually watching a small portion of the video
The amount of data gets downloaded
IT makes sense to utilise this technique to reduce the amount of data that you are downloading.

Content distribution network
It is in between our device and the cloud server
Sort of mini cloud.
It can cache that data in the CDN

Web Page Load

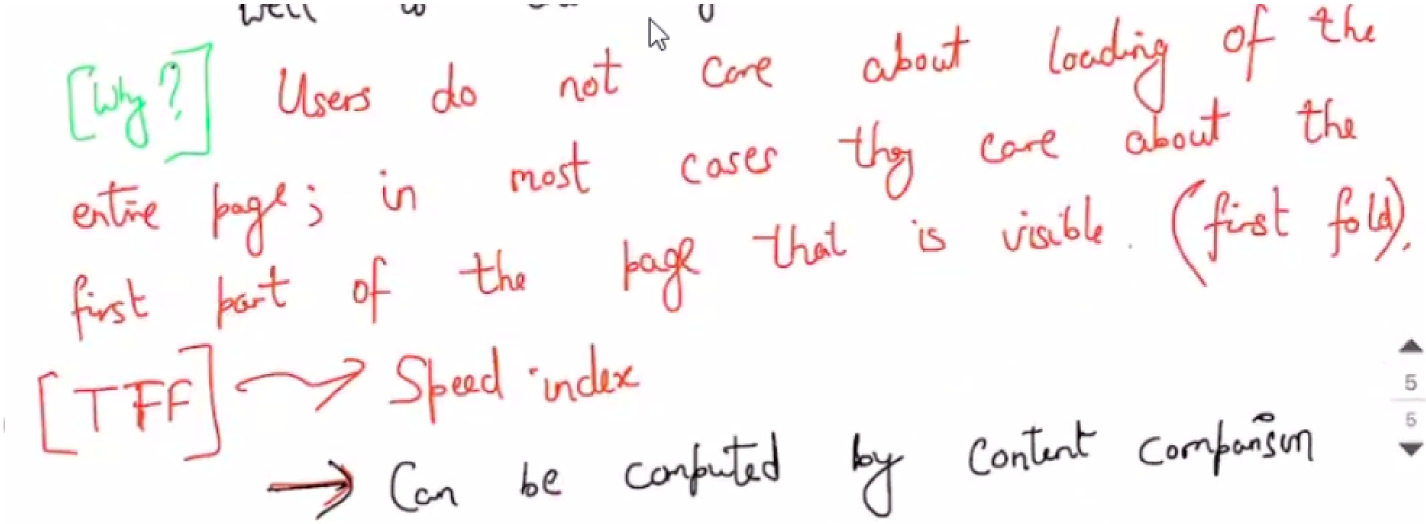
We talked of QOE ex throughput
We talk about the page load time, PLT
This is the easiest metric to measure simply because of the way this browser works
How HTTPS/HTTP work

Ususally we have something calle JS in our browsers,
When the page loads, then we have an on load function, and it executes when the entire page loads.
By monitoring the time taken to execute this on load function. We can find out the page load time.

Here, some challenges might come in
If we want to pptimise we can always do so using variety of techniques suppose we try to ensure that even in web pages, we have dependent objects, based on the priority.
It has some complications, this page load time is not a very good QOE metric.(Objective)
It does not map very well to the subjective score.
Why?

Users do not care loading about the entire page.
In most cases, they only care about loading the first part of the page which is visible.

Time to first fold.(TFF)
Can be computed by content comparison



TFF is also objective.
On load function executed at the end of the page load.
IT does not really explain, the users care about some part of the content, not all part of the content.
Users care about the content.

Using some 20-30 users in the room and doing this gaze tracking.
Using some controlled experiments it is possiBLe to identify the location where the user will focus.

Based on this, we can assign priority to each different object.

If the website is not optimized, then ads come fast, ads use CDN extensively.
Whatever the user cares about, it is called the User page load time.
We can maximize the UPLT given some specific network condition.

Assumption was QoE is a linear metric. So far, our assumption was that QoE is a linear metric.
Somehow based on user perception.
In practice QoE is a logarithmic function.

Why logarithmic
Whether we consider the bit rate, or time or something else, but it turns out that human perception is logarithmic in nature.
If we increase the resolution from 240P to 360P

Perception is we can see a very large improvement.
360P to 480P there is a smaller improvement.

This thing is there for this rebuffering time also.
If the rebuffering increases, the amount of dissatisfaction is larger.
Opinion score.
1 to 1.5s

practice, QoE is a logari

240p to 360p $\rightarrow$ immediatly see a large improvemen

360p to 480p $\rightarrow$ see a smaller improvement than above.

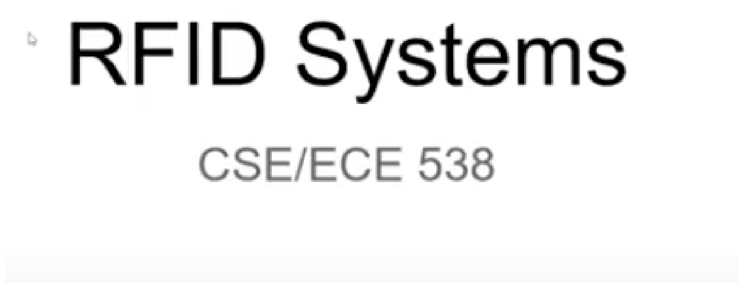
Rebuffing time $\rightarrow$

0.5 $\rightarrow$ 1s } $\rightarrow$ Your dissatisfaction level ~~is the~~ increases by a higher amount

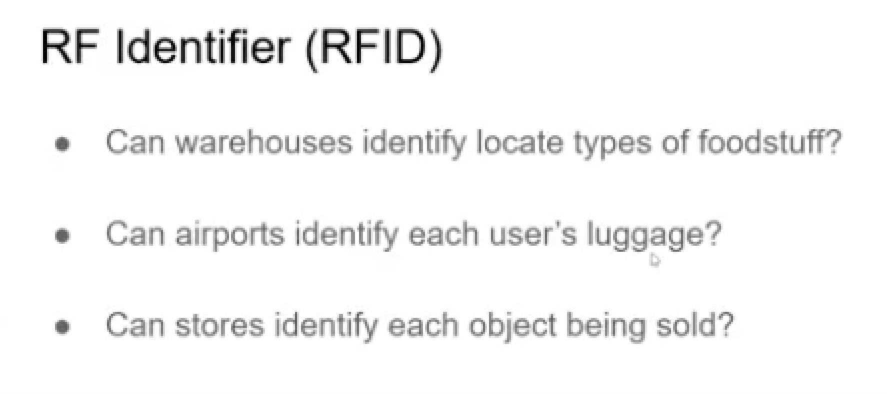
1s $\rightarrow$ 1.5s

QoE uses the logarithm

Lecture 22

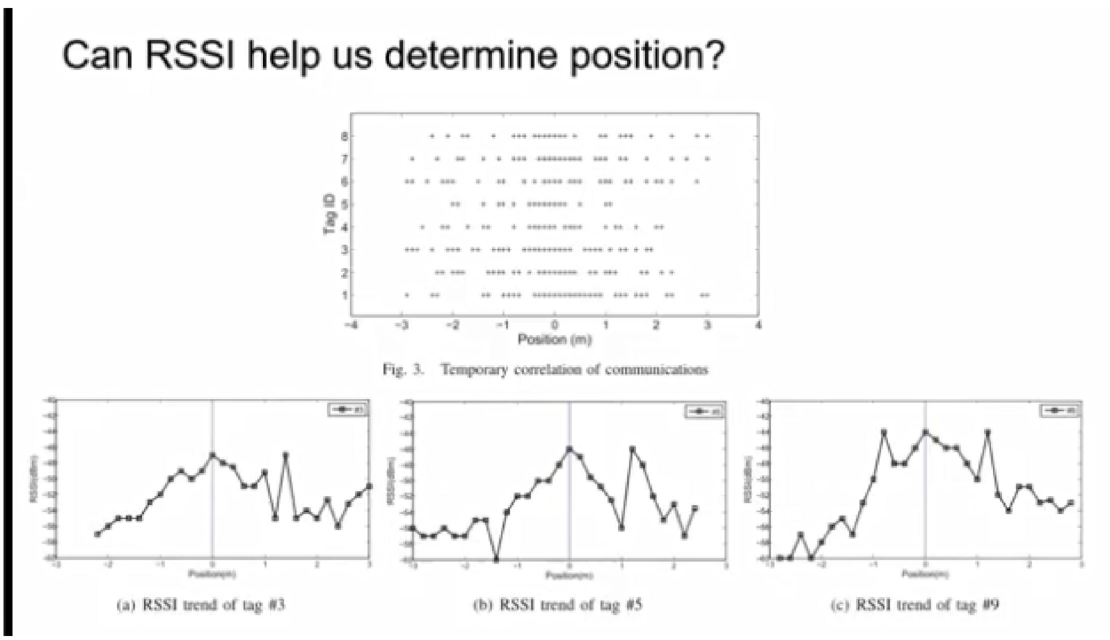


We have discussed outdoor localisation using GPS via time of flight.
Whenever we talk about localisation, we talked about people, etc.
Bar codes are there in the back of book, the bar code encodes some information and using that bar code,we can see what information is there. RFID stands for radio frequency of Identification, It's range in terms of distance is larger. So it is beter than barcodes. Range of RFID is larger in terms of distance, you can still do it. RFID encodes information efficiently than barcode. Barcode takes larger area as it is encoded in spacial domain, RFID are encoded in TIME domain. So, time domain implies, across time, all data is bits, depending on the sequence of bits encoded, either the signal goes back or it doesn't. Third one is RFID not only encides the information but also be able to localise.RFIS;s go to cun centimeter level)
There is nothing very interesting in that. But RFID's have this ability to identify where the object is. The technolog is far more accurate in localisation than GPS etc. They all provide 10cm level accuracy. On GPS iff you get 6 to 7 satellites. In most cases we get metres level. RFID goes to sub centimeteres level.



Currently they are used for food, user's luggage etc.
Airport attach RFID to each luggage so that it does not get misplaced.
RFID vs BARCODE:
RFID have much longer range than BARCODE readers.
Passive RFID has no energy source on the tag side. Energy source is only on the reader side. Active RFID has batteries. So that it increases the cost. Passive one is Very cheap in practice. Backscatter communication.

Transmitter on the reader side and tag side is receiver side. Tag is dark in color so that they absorb light, so there is some amount of current generated in receiver antenna. The tag side ahs something related to impedance matching. The impedence is of the TAG's internal circuit. It can be changed in a specific pattern to indicate the encoded data. This is known as impedance matching. If we want to send 0, then your TAG can increase the impedance. Impedance reduces the signal power. High impedance reduces the signal power. The signal that is emitted from the receiver side will be lower. Data of 0's and 1's across time can be sent by changing the impedance accordingly. That is why this is in time domain. There is no spacial component to this. The transmitter is sending the power and that power is used to generate the singal on the receiver side, the tag does not need a battery. Apart from this, if we have to change the data, we need to paste it on the already existing bar code. It is ROM. We have EEPROM.

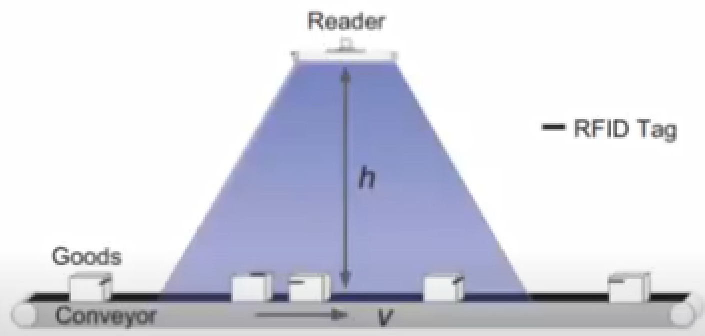


Challenge is to identify which parameter gives us the required location.
RRR of ith item is number of responses received from i in d / number of expected responses from i in d.
This alone does not provide sufficient clarity.
There are three components of order tracking for luggage in mobile based RFID systems

1. Moving conveyer belt with a velocity v
2. A RFID reader fixed over the belt at a height of h
3. A sequence of luggage attached with RFID tags on the belt.

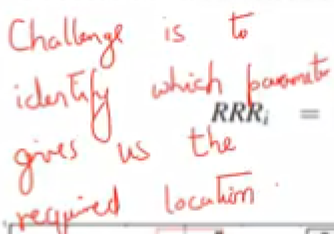
Tags will be accessed multiple times during the movement along the belt.
The only information that we can use to track the order of tags is the received responses when the tags are within the communication rage of the reader. We know that due to the randomness from tags' are within the communication range of the reader. We know that due to the randomness from tags' replies, the communications by themselves do not contain any clue to infer tags' relative positions.

OTrack: Order Tracking for Luggage in Mobile RFID Systems

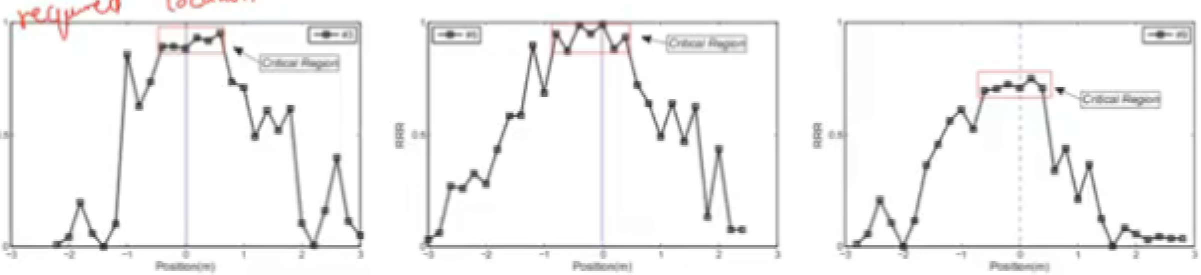


RSSI is received signal strength indicator.
This part of wireless network is close to data science. Wireless network people are doing all sort of things. The challenge is to identify that which parameter gives us the required location. RSSI does not directly

What is the alternative?



$$RRR_i = \frac{\#\text{ of responses received from } i \text{ in } d}{\#\text{ of expected responses from } i \text{ in } d},$$



(a) RRR trend of tag #3 (b) RRR trend of tag #5 (c) RRR trend of tag #9

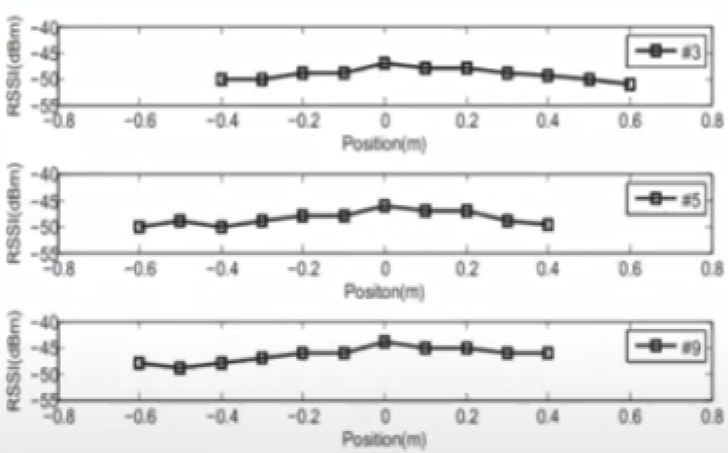


Fig. 6. Combination of RSSI and RRR

Within this critical region, look for the RSSI
RSSI works well when we have less ambiguity.

This is a principle of RSSI based localisation. It can work well with relatively close range, if you go away range gets worsened. Within that approximate region

TagORam

Here we use the phase difference. Instead of using RSSI, we can use the concept of phase difference, If the object is at d distance, the signal has to travel 2d distance. The phase difference theta hat will be

$$\tilde{\theta} = \left(\frac{2\pi}{\lambda} \times 2d\right) \bmod 2\pi$$

1) Transmission by the reader
2) Reception by the tag
3) Transmission by the tag

$\left.\begin{array}{l}\text{2) Reception by the tag}\\ \text{3) Transmission by the tag}\end{array}\right\}\theta_{TAG}$

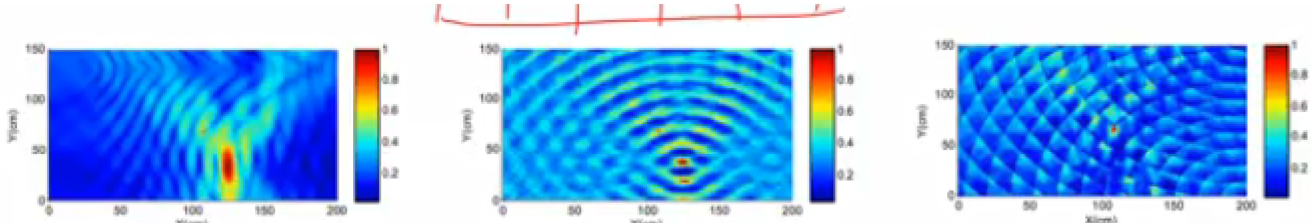
4) Reception by the reader

$$C = \theta_T + \theta_{TAG} + \theta_R$$

Also some amount of noise added.

The noise is gaussian in nature, but c is known as diversity term and it affects the final localisation result. Each of the TAG has distinct amount of phase difference that comes in. The actual phase is not a property of distance anymore, it depends on type of tag, tag number and also on the distance. Calibrating the theta tag is more challenging. Divide the entire area into the grid. Assume that there is no diversity, Assume that there is only this noise factor.

We already have probability gaussian distribution. Unfortunately, we will get something like one red and all are blue



There is a Gaussian distribution of the phase values.

How can you calibrate this to identify as the object is moving, instead of one particular datapoint. We have multiple antennas and multiple readings.
Each theta ij represents the phase measurement by antenna i for the jth time.
Time intervals might be different.

$$T = \begin{bmatrix} T_{11} & \dots \dots, & T_{1m} \\ & & \vdots \\ & & T_{nm} \end{bmatrix}$$

Assumption: $\rightarrow$ Velocity of the objects is constant and known.

$$f(t_k) = f(t_0) + \vec{V} \cdot \Delta t \qquad \Delta t = T_{ij} - T_{11}$$

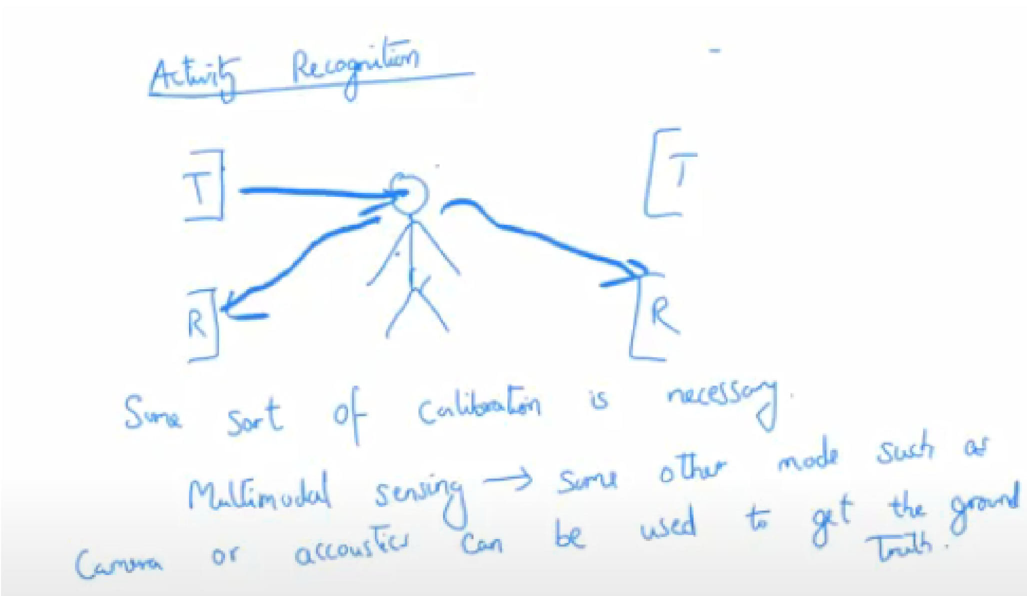
We can form a system of equations and then perform a regression with the data received.
We have noise in all wireless systems. What can we do to reduce the impact of noise?

Form a system of equations and then perform a regression with the data received.

The regression error will reduce.

Activity Recognition

Suppose you have a series of transmitters and some receivers in a room and whenever someone is in the room and there is some activity happening. Each of the activity actually reflects signals in slightly different way. Receiver identify how is this thing changes. We can do some sort of calibration. Even this calibration is relatively easy now. There is something known as multi-modal sensing. It is used in computer vision, Either using a camera or using the sound that the user makes, we can identify what is the activity that they are doing and that can train your system. We are tying to find the ground truth. This multi modal sensing implies that some other mode



If you can the CSI in the form of amplitude and phase can serve as a signature for each activity. Sign is not very noisy, it can happen that these signatures are generally noisy in nature. We can encode them in terms of time series and use it as a classification problem adn the classes are each different type of activities. Barnet used deep neural network to classify the time series. But other techniques are also possible.

Lecture 23

TCP → How will TCP respond if something like this happens

16. Bandwidth increases or falls
17. Latency increases or falls
18. With variant performs better under which conditions

Energy Efficiency → Focus on wireless networks and smartphones

1. NAPMAN
2. MICROSLEEP
3. What is the advantage and disadvantage of these techniques
4. Which components can consume additional energy
5. Tail energy in cellular networks
6. If we get the energy consumption from a time perspective then we will see that as we are using better and better cellular network protocols, energy consumption per unit time is increasing. But energy consumption per unit bit is decreasing
7. Our cellular protocols are encouraging higher consumption of the data because if you conserve data that means you are losing out, phone's efficiency is lower then
8. Numerical problems related to energy consumption
9. Energy consumption of the display system

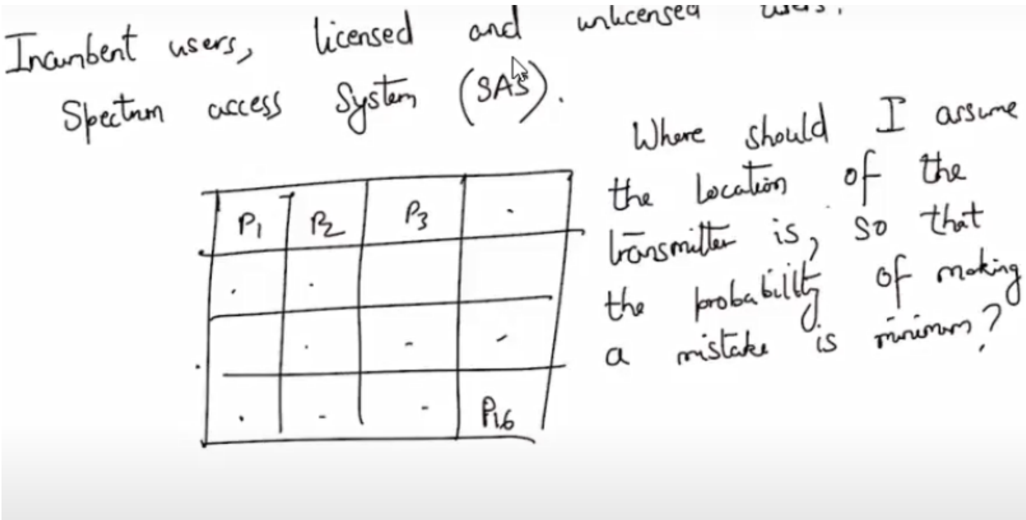
NS-3

1. Some questions about simulation
2. Advantage of simulation using NS3 vs other evaluation techniques like test bed.
    a. Simulation environment has more flexibility in terms of how you evaluate, test bed we cannot change hardware
    b. Simulation in NS3 is considered to be more realistic, the modules have already been coded as libraries

GPS

1. GPS uses CDMA technology. The communication part uses CDMA, the localisation and time synchronization part uses this time difference of arrival.
2. Suppose we have 3 satellites instead of 4 satellites, what will happen if there is a drift in the clock?
3. There is a concept of active antenna, suppose when we come indoors, GPS does not work, outdoors also the power is low, how will GPS work? Antenna has amplifier
4. In GPS, one issue is , because the satellites are so far away, and they do not have lot of energy available, the power is relatively high, by the time it reaches the earth, it is quite low, at that power level, it is below and the antenna senses through the power, the power is so loqw that the signal gets lost. Antenna itself has an attached amplifier 2. We call a passive antenna, the standard antenna available for cell towers, are passive antennas. They do not need any amplification, you attach it into anything, they will start working. GPS antennas require connection to a power source for them to work.

Dynamic Spectrum Access

1. Wireless Sensor networks
2. Concept of localisation by using fingerprints
3. Privacy aspects of dynamic spectrum access
4. Citizen broadband service
5. CellFi uses cellular network protocol over spectrum holes. Wifi is not good for long range
6. Incumbent users, licenced and unlicenced users
7. Spectrum access System(SAS)
8. 

9. The probability is maximum, there we will keep the transmitter.
10. Suppose from prior knowledge that maybe one of the cells is going through the load, suppose the transmitter is at that load, multiply it with the prior, then normalize it. Consider prior probabilities if they are not uniform.
11. Selecting sensors to improve efficiency of spectrum access.
12. Fusion rule → Combining the results of multiple sensors
13. Approximation ratio, what is in polynomial time and what is in non polynomial time
14. Monotone and submodular, follow the greedy algorithm
15. Heuristic and approximation algorithm
16. The approximation algorithm is worse than getting an optimal algo. Optimal algo is nothing but an exhaustive search. There is nothing better than that. Heuristic does not give a guarantee. Maybe you can find that using some heuristic. Greedy will come under approximation.
17. Suppose you have a few sensors like this. Now you have decided that your budget is such that you can select.

Quality of Service and Quality of experience

Video streaming and web browsing and bitrate, rebuffering, and changes of resolution all affect the quality
Formulated as an optimisation problem and solved user perception measured via mean opinion score.
Quality of service is network throughput → Jitter, latency etc.
Web browsing QOE → Page load time is not a good metric.
If the network is not so good, you can still do lot of thing, adapt to the network, buffer as much as you can at the application layer. You have page load time, which is not a good metric. Why?

What are the other metrics?
Speed index, user page load time (UPLT)
If you do not scroll down, how much time do you take?
Focus on the parts of the webpage that users focus on.

RFID tags : Focus was on two works: RSSI Based localisation of object.
Otrack and TagoRAM
One is RSSI based localisation and other is phase difference based localisation.
RSSI and RRR(response rate) in otrack
TagORam is more complex data filtering technique.

RRR gives more accurate course grain idea of the location.
RSSI does not work at course grain, but works well when the location is somewhat accurate via RRR.

RSSI based localisation is sensitive to reflecting surfaces
The moment object is coming closer to the reflecting surface, then the signal is coming direct

We focus mostly on the passive parts. The devices do not have any battery. The tags have battery,.
This is passive backscatter technique with no energy present in batteries.

LTE and WIFI coexistence.