1 Introduction

Most networks are organized as a stack of layers in a hierarchical manner, each one built upon the one below it. The functions of layer, the contents of the layer, number of layers differ from network to network. The purpose of each layer is to giving certain services to the higher layers preventing those layers from the details of how the offered services are actually implemented.

Various Layers of design starting from lower level

1) Physical Layer
2) Data link layer
3) Network layer
4) Transport layer
5) Application layer

Entities at each layer communicate with each other using appropriate protocol. A protocol relates to the implementation of the service and is not visible to the user of the service. Lower layers provides certain services to higher layers. A service is set of operations that layer provides to the layer above it. The service does not say how the operations are implemented. Layers can offer two different types of services to the layers above them: Connection-oriented and connectionless service. Lower layer hide details from the upper layers. The actual communication and data transfer takes place at the lower layer which is the physical layer.

Physical Layer

It is responsible for transmitting stream of bits over a communication channel. In wired system bits are transferred on a wire. But on wireless systems there are various issues like choosing carrier frequency for modulation or detecting carrier signals. The wireless communication is unreliable in physical layer because of channel noise and interference.

Data link Layer

In data link layer here we are talking about mainly MAC layer. The various functions of MAC layer are

a) accessing the medium
b) multiplexing different data streams
c) correction of transmission errors
d) prevent collision among various stations who are transmitting

The goals of MAC layer for wireless communication is to achieve higher level of reliability and reasonable throughput. Broadcast networks helps control access to shared channel with the help of MAC layer.
2 MAC layer strategy for wired network:

CSMA/CD - Carrier sense multiple access/collision detection
This is used mainly in Ethernet. It is a channel sharing strategy for sharing wired channels. Here sender senses medium whether it is idle or not.
- If medium is busy then sender waits until channel becomes free.
- If the channel is free then it transmits data but continues to listen to the channel. At any point if a collision occurs then sender aborts transmitting.
This means that if two stations sense the channel to be idle and began transmitting they will both detect the collision almost immediately. Rather than finish transmitting they should stop as soon as collision is detected. This helps in saving time and bandwidth.

3 Wireless Network

If sender detects collision it does not help as there might be deterioration of signal strength between sender and receiver. As distance increases the distance travelled by signal also increases which results in signal strength going down. This is due to the fact that strength of signal is inversely proportional to the square of distance.
No collision at sender does not imply that there is no collision at receiver.
This can be illustrated by taking some special cases which are mentioned as following.

3.1 Hidden terminals

Suppose there are three stations A,B and C. In Figure 1 (page 6), A is within B’s transmission range and B is also within C’s transmission range. But A and C station cannot hear each other as C is not within A’s range. A wants to send to B and C wants to send to B. Station C starts transmitting to station B. If A senses the channel it will not hear anything and falsely concludes that it may now start transmitting to B. In such a scenario collision would happen as both C and A are sending to B. So this will cause an undetected collision at B. This is a problem for the medium because the competitor is too far away which is called as hidden station problem. The solution to this problem can be that the detection range can be made larger than the reception range and this can help to avoid the above problem.

3.2 Exposed Terminals

In Figure 2 (page 7), B is within A’s and C’s transmission range, C is within B’s and D’s transmission range. B and D cannot hear each other as D is not within B’s transmission range. Suppose if B starts transmitting to A. If C senses the medium it will hear an ongoing transmission from B and falsely conclude that it may not send to D because B’s and C’s transmission may collide at D. But in fact such a transmission is possible as B’s transmission range does not include D. So this is exactly opposite to the previous hidden terminals case. A problem is detected but actually it does not exist. This is called as Exposed Terminals problem.
3.3 Near/Far Problem

We take an example to show Near/Far Problem. Suppose there are 3 stations A, B, C in a line. B is between A and C stations. A wants to send to C and B also wants to send to C. The distance of A is more from C then B due to there topology. So while going to C from A the signals have to pass through B which causes degradation of A’s signal when trying to reach C.

This is caused because B’s signal drowns A’s signal, which has already been weakened due to distance from A, when both are sending with same power frequency.

4 Detecting collision at receiver

There are mechanisms for detecting collision at Receiver end.

4.1 Acknowledgements (ack)

Receiver sends an ack for each data message. But it also gives rise to one sided error.

If ACK received by sender then message received successfully by receiver.
If ACK not received by sender this does not mean that message was not successfully transmitted. There could have been the case when ack was lost due to noise.

4.2 Reservation of slots

(MACAW by Bhargavan, Demens et al SIGCOMM ’94)

Data can be very long so handshake is preceded before actual data can be send. This can be done by sending short control messages. Handshake reserves the channel for a certain period. This is a kind of virtual channel sensing. This is done by RTS/CTS/Data frames.

RTS- Request to send data
CTS-Clear to send data

The various steps involved are as follows:

a) Sender A sends RTS to receiver B. This contains short header (names of sender and receiver and length of data to be send).
b) Receiver B replies with a CTS with the same matching information.
c) When sender A receives CTS then it starts transmitting the data.

This sequence can be interrupted if sender hears another different CTS then it is expecting. Also if receiver B hears another RTS within the same time from someone other than A then B does not respond with CTS because otherwise collision might occur.

Advantages
i) This method does not require carrier sensing.
ii) It uses explicit messages.

iii) It reduces collision. RTS/CTS may collide but the probability is less as the messages are very short.

iv) It solves the hidden terminal problem. A sends RTS to B. C sends RTS to B. B replies with a CTS only to A. Here A and C cannot hear to each other.

v) It solves exposed terminal problem. B sends RTS to A. C sends RTS to D. A responds by sending CTS to B. D responds by sending CTS to C.

This method is most useful when the data packets are long and collisions are frequent but if data packets are short and collisions are infrequent then the overheads would be much larger than in the previous case.

The important point to note here is as the transceivers are half duplex they cannot listen and transmit at the same time so the problem of interference of RTS or CTS signals is less likely to happen.

Another variation of previous method is including acknowledgements (RTS/CTS/Data/ACK). It makes sure that even if the data packets are long and collisions are frequent then such method can be helpful in reducing probability of collision.

5 CSMA with Collision Avoidance

In this protocol, time slots are divided for transmission and in each slot we can transmit with certain probability of transmission. If a collision occurs, the colliding stations wait for next time slot using the binary backoff algorithm and try again later to send data. Assume slots to be of equal length which is equal to frame size. This method is good only for unicast systems.

There are two cases, one where stations transmit at slot boundaries, and the other where stations can transmit at any time within the slot.

a) All stations transmit at slot boundaries. This means they start to transmit at the beginning of the slot. This is a slotted aloha concept.

In this case, users need to transmit at slot boundaries. So, time synchronization is an issue. We would be talking about how to handle this in later lectures.

Set up probability \( P \) for transmission at the beginning of the time slot. What is the probability that host \( i \) successfully transmits in a particular time slot?

\[
P(1 - P)^{N-1}
\]

Here \( P \) is the probability of transmission of host \( i \) in the time slot and \( N \) is the number of stations. \( 1 - P \) is the probability of each station not transmitting, so \( (1 - P)^N \) is the probability that all other \( N - 1 \) stations do not transmit in that time slot.

So, the expected number of successful transmissions = \( NP(1 - P)^{N-1} \).
This is maximized at $P = 1/N$, (As we maximize here so this gives the best bound for transmission in that time slot), giving

$$(1 - 1/N)^{N-1}$$

As $N$ gets larger,

$$\lim_{N \to \infty} (1 - 1/N)^{N-1} = 1/e$$

b) Stations can transmit at any time, but once within each time interval. This is pure aloha. In this case users still transmit at local time slots, but there is no synchronization between these slots. Hence, no time synchronization is needed between stations.

In Pure aloha, suppose a station starts to transmit at time $t$. Since the packet will arrive by time $t+1$, no other packet must be enroute during time interval $[t, t+1]$, to avoid collisions. A packet whose transmission starts during time $[t-1, t+1]$ would still be enroute at time $[t, t+1]$. So, two time slots need to be avoided to prevent collision.

For successful transmission, the probability would therefore become half.

$$P(1 - P)^{2(N-1)}$$

$(1 - P)^{2(N-1)}$ is the probability of all other N-1 stations not transmitting for 2 time slots to prevent collision.

Expected number of successful transmissions $= NP(1 - P)^{2(N-1)}$. This is maximized at $P = 1/N$, giving

$$(1 - 1/N)^{2(N-1)}$$

As $N$ tends to infinity, we get

$$\lim_{N \to \infty} (1 - 1/N)^{2(N-1)} = 1/2e$$

5.1 BackOff Algorithm

If a collision occurs then we want to wait longer for retransmission. Let $BO =$ Largest number of slots to wait.

Pick backoff [0,BO] at random.

Initialize counter.

Decrement counter at each idle interval.

When counter reaches zero then transmit.

In the class one question arose that what would happen if two stations wait for the same amount of time before transmitting? The solution could be if a collision occurs and two stations wait for the same amount of time before transmitting. This could result in another collision therefore it is necessary to wait for random backoff so so that possibility of collisions is decreased.
One advantage of this backoff algorithm is that it controls the medium load. If the medium is heavily loaded, the likelihood of collisions increases, and the algorithm increases the interval from which the random delay time is chosen. This should lighten the load and reduce further collisions.

By having the random interval grow exponentially as more collisions occur the algorithm ensures low delay when only a few stations collide but also ensures that collision is resolved in a reasonable time interval when there are more number of collisions.

This approach causes a station which has had a collision to be more likely to be delayed again which may be unfair. An alternative approach could be to decrease Backoff if a collision happens. This reduces the likelihood of a station that has had a collision to have further delays.

6 Figures

![Diagram](image)

Figure 1: A and C are hidden from each other as A is within B’s range and C is within B’s range but A and C are not within each other’s range.
Figure 2: A is within B’s range, B is within C’s range, C is within D’s range