Chapter 4. The Medium Access Control Sublayer

The Medium Access Control Sublayer

- Networks can be divided into two categories: those using point to-point connections and those using broadcast channels.
- In any broadcast network, the key issue is how to determine who gets to use the channel when there is competition for it.
- When only a single channel is available, determining who should go next is much harder.
- Broadcast channels are sometimes referred to as **multiaccess channels** or **random access channels**.

- The protocols used to determine who goes next on a multiaccess channel belong to a sublayer of the data link layer called the **MAC** (**Medium Access Control**) sublayer.
- The MAC sublayer is especially important in LANs, many of which use a multiaccess channel as the basis for communication.
- WANs, in contrast, use point-to-point links, except for satellite networks.
- Technically, the MAC sublayer is the bottom part of the data link layer.

4.1 The Channel Allocation Problem

- **Static Channel Allocation (Fixed Channel allocation)**
- **Dynamic Channel Allocation**

<u>Static Channel Allocation in LANs and MANs</u>

- The traditional way of allocating a single channel, such as a telephone trunk, among multiple competing users is Frequency Division Multiplexing (FDM).
- If there are *N users, the bandwidth* is divided into *N equal-sized portions, each user being assigned one portion.*
- Since each user has a private frequency band, there is no interference between users.
- When there is only a small and constant number of users, FDM is a simple and efficient allocation mechanism.

- **Disadvantage:** when the number of senders is large and continuously varying, FDM presents some problems.
- If the spectrum is cut up into *N regions and fewer than N users are currently interested in communicating, a large piece of valuable spectrum will be* wasted.
- If more than *N users want to communicate, some of them will be denied permission* for lack of bandwidth, even if some of the users who have been assigned a frequency band hardly ever transmit or receive anything.
- In TDM based allocation Each user is statically allocated every *Nth time slot. If a user does not use the allocated* slot, it is just wastage of bandwidth.

Dynamic Channel Allocation in LANs and MANs

- In dynamic channel allocation scheme, frequency bands are not permanently assigned to the users.
- Instead channels are allotted to users dynamically as needed.
- This allocation scheme optimizes bandwidth usage and results is faster transmissions.
- The allocation is done considering a number of parameters so that transmission interference is minimized.

- All the work done in this area are based on five key assumptions:
- **1. Station Model:** The model consists of *N independent stations (e.g., computers, telephones,* or personal communicators), each with a program or user that generates frames for transmission. Stations are sometimes called **terminals.**
- Once a frame has been generated, the station is blocked and does nothing until the frame has been successfully transmitted.

- **2. Single Channel Assumption:** A single channel is available for all communication. All stations can transmit on it and all can receive from it.
- **3. Collision Assumption:** If two frames are transmitted simultaneously, they overlap in time and the resulting signal is garbled. This event is called a **collision.** All stations can detect collisions.
- A collided frame must be transmitted again later. There are no errors other than those generated by collisions.

- **4a. Continuous Time:** Frame transmission can begin at any instant. There is no master clock dividing time into discrete intervals.
- **4b. Slotted Time:** Time is divided into discrete intervals (slots). Frame transmissions always begin at the start of a slot.
- **5a. Carrier Sense:** Stations can tell if the channel is in use before trying to use it. If the channel is sensed as busy, no station will attempt to use it until it goes idle.
- **5b. No Carrier Sense.** Stations cannot sense the channel before trying to use it. They just go ahead and transmit. Only later can they determine whether the transmission was successful.

Difference

	FCA	DCA
Channel Allocation	Fixed number of channels or voice channels are allotted.	Channels to be allotted are not fixed initially.
Blockage	If all channels are occupied, then user call is blocked.	If all channels are blocked, then Base Station(BS) requests more channels from Mobile Station Center(MSC).
Algorithm	No need to complex algorithm.	Algorithm to determine efficient channel availability is quite complex in DCA.
Cost	FCA is cheaper than DCA.	DCA is costly as real time computation needed.
Cell Allocation	Once call is complete, channel remains with the cell.	Once call is complete, channel is returned back to Mobile Station Center.

4.2 Multiple Access Protocols

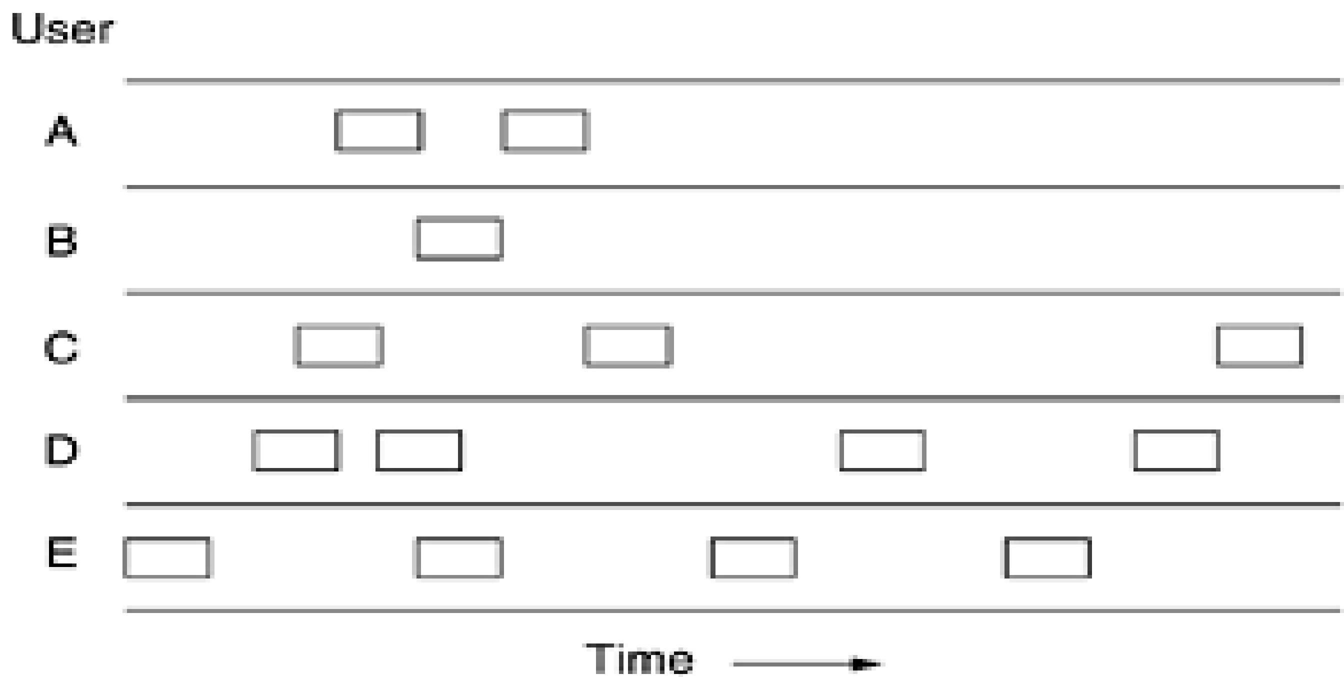
ALOHA

- In the 1970s, Norman Abramson developed this method to solve the channel allocation problem.
- Two versions of ALOHA here: **pure and slotted**.
- They differ with respect to whether time is divided into discrete slots into which all frames must fit.
- Pure ALOHA does not require global time synchronization; slotted ALOHA does.

Pure ALOHA

- The basic idea of an ALOHA system: Users transmit whenever they have data to be sent. There will be collisions, of course, and the colliding frames will be damaged.
- However, due to the feedback property of broadcasting, a sender can always find out whether its frame was destroyed by listening to the channel.
- With a LAN, the feedback is immediate; with a satellite, there is a delay of 270 msec before the sender knows if the transmission was successful.
- If listening while transmitting is not possible for some reason, acknowledgements are needed.
- If the frame was destroyed, the sender just waits a random amount of time and sends it again.
- Systems in which multiple users share a common channel in a way that can lead to conflicts are widely known as **contention** systems.

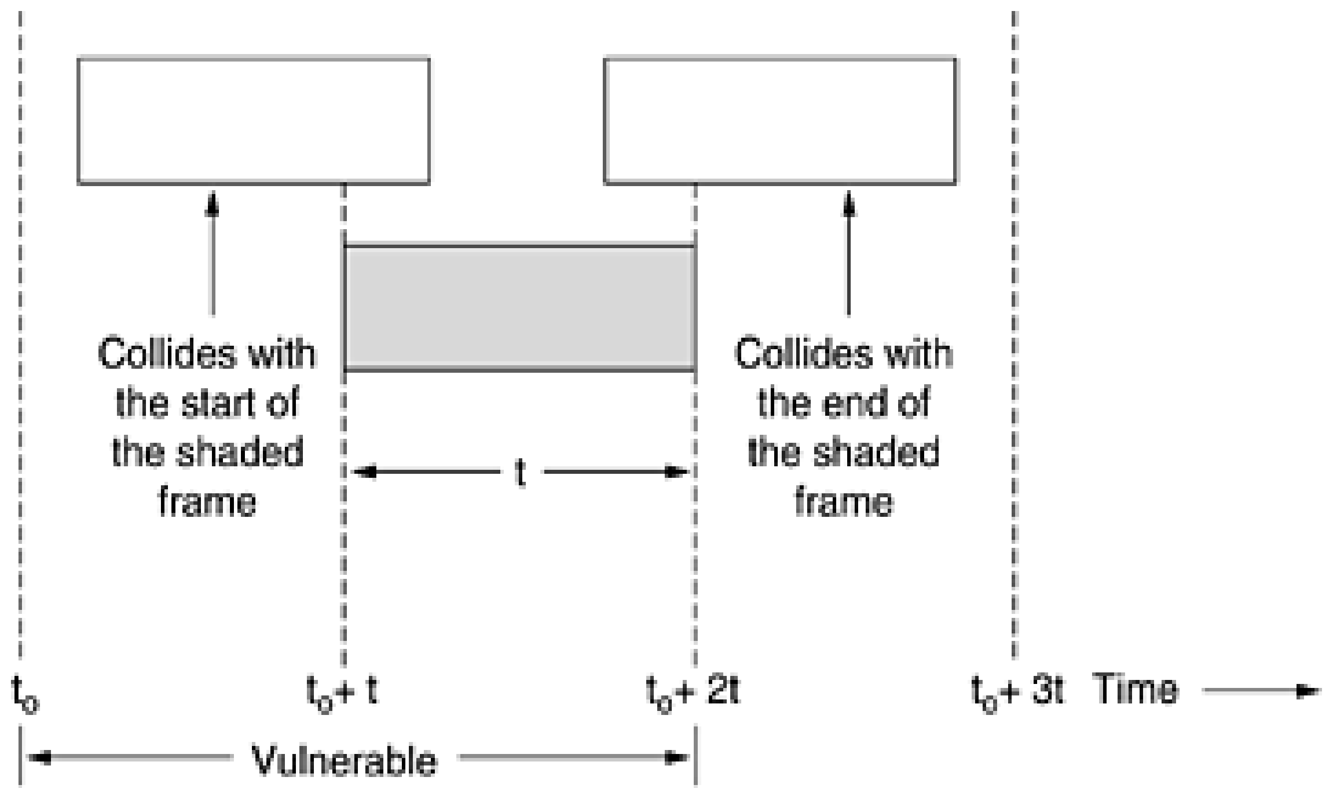
Figure 4-1. In pure ALOHA, frames are transmitted at completely arbitrary times.



- Whenever two frames try to occupy the channel at the same time, there will be a collision and both will be garbled.
- The checksum cannot (and should not) distinguish between a total loss and a near miss.
- A user is always in one of two states: **typing or waiting.**
- Initially, all users are in the typing state. When a line is finished, the user stops typing, waiting for a response.

- The station then transmits a frame containing the line and checks the channel to see if it was successful. If so, the user sees the reply and goes back to typing.
- If not, the user continues to wait and the frame is retransmitted over and over until it has been successfully sent.
- Let the "**frame time**" denote the amount of time needed to transmit the standard, fixed-length frame (i.e., the **frame length divided by the bit rate**).

Figure 4-2. Vulnerable period for the shaded frame.



- **A frame will not suffer a collision if no other frames are sent within one frame time of its start.**
- Under what conditions will the shaded frame arrive undamaged?
- Let t be the time required to send a frame. If any other user has generated a frame between time $t0$ and $t0 + t$, the end of that frame will collide with the beginning of the shaded one.
- **In pure ALOHA a station does not listen to the channel before transmitting, it has no way of knowing that another frame was already underway.**
- Similarly, any other frame started between $t0 + t$ and $t0 + 2t$ will bump into the end of the shaded frame.
- An interesting question is: What is the **efficiency of an ALOHA channel?** what fraction of all transmitted frames escape collisions under these chaotic circumstances?

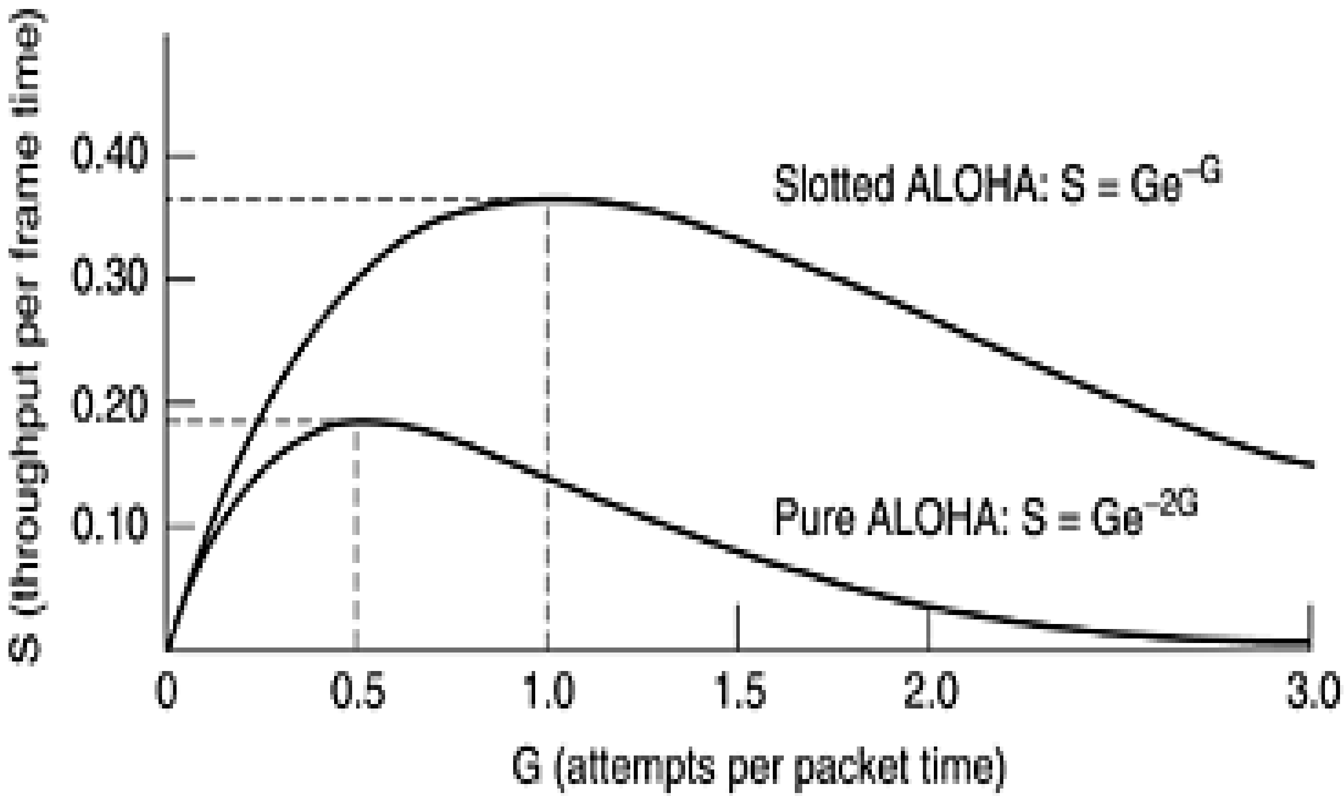
- The probability that *k* frames are generated during a given frame time is given by the Poisson distribution:

$$\Pr[k] = \frac{G^k e^{-G}}{k!}$$

- so the probability of zero frames is just e^{-G}.
- Throughput $S = Ge^{-2G}$

- The relation between the offered traffic and the throughput is shown in Fig. 4-3.
- The **maximum throughput** occurs at $G = 0.5$, with $S = 1/2e$, which is about **0.184**.
- In other words, the best we can hope for is a channel utilization of 18 percent.

Figure 4-3. Throughput versus offered traffic for ALOHA systems



Slotted ALOHA

- In **1972, Roberts** published a method for doubling the capacity of an ALOHA system.
- **His proposal was to divide time into discrete intervals, each interval corresponding to one frame.**
- In Roberts' method, which has come to be known as **slotted ALOHA**, in contrast to Abramson's **pure ALOHA**, **a computer is not permitted to send whenever a carriage return is typed.**
- Instead, **it wait for the beginning of the next slot.** Thus, the continuous pure ALOHA is turned into a discrete one.
- Since the vulnerable period is now halved, the probability of no other traffic during the same slot as our test frame is *e-G* which leads to

$$S = Ge^{-G}$$

- **Slotted ALOHA peaks at *G* = 1, with a throughput of *S* =1/*e* or about 0.368, twice that of pure ALOHA.**
- If the system is operating at *G* = 1, the probability of an empty slot is 0.368.
- **The best we can hope for using slotted ALOHA is 37 percent of the slots empty, 37 percent successes, and 26 percent collisions.**
- Operating at higher values of *G* reduces the number of empties but increases the number of collisions exponentially.
- When Internet access over the cable was invented, all of a sudden there was a problem of how to allocate a shared channel among multiple competing users.

Carrier Sense Multiple Access Protocols

Carrier Sense Multiple Access Protocols

- Carrier Sense Multiple Access (CSMA) is a network protocol for carrier transmission that operates in the Medium Access Control (MAC) layer.
- With slotted ALOHA the best channel utilization that can be achieved is $1/e$.
- In local area networks, however, it is possible for stations to detect what other stations are doing, and adapt their behavior accordingly. These networks can achieve a much better utilization than $1/e$.
- Protocols in which stations listen for a carrier (i.e., a transmission) and act accordingly are called **carrier sense protocols**.
- It senses or listens whether the shared channel for transmission is busy or not, and transmits if the channel is not busy.
- Using CSMA protocols, more than one users or nodes send and receive data through a shared medium that may be a single cable or optical fiber connecting multiple nodes.

Working Principle

- When a station has frames to transmit, it attempts to detect presence of the carrier signal from the other nodes connected to the shared channel.
- If a carrier signal is detected, it implies that a transmission is in progress.
- The station waits till the ongoing transmission executes to completion, and then initiates its own transmission.
- Generally, transmissions by the node are received by all other nodes connected to the channel.
- Since, the nodes detect for a transmission before sending their own frames, collision of frames is reduced.
- However, if two nodes detect an idle channel at the same time, they may simultaneously initiate transmission. This would cause the frames to garble resulting in a collision.

Persistent and Nonpersistent CSMA

1-persistent CSMA

- The first carrier sense protocol is called **1-persistent CSMA** (Carrier Sense Multiple Access).
- 1-persistent CSMA is a Carrier Sense Multiple Access (CMSA) protocol that operates in the Medium Access Control (MAC) layer.
- In 1-persistent CSMA, when a transmitting station has a frame to send and it senses a busy channel, it waits for the end of the transmission, and transmits immediately.
- Since, it sends with a probability 1, the name 1 – persistent CSMA is given.

Algorithm

The algorithm of 1-persistent CSMA is:

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is busy, the station waits and continually checks until the channel becomes idle.
- If the channel is idle then it transmits the frame immediately, with a probability 1.
- A collision may occur if two or more channels transmit simultaneously.
- If collision occurs, the station waits for a random period of time and restarts the algorithm all over again.

Disadvantages of 1-persistent CSMA

There are chances of collisions in the following situations:

- The propagation delay has an important effect on the performance of the protocol.
- The longer the propagation delay, the more important this effect becomes, and the worse the performance of the protocol.
- **Situation 1:** Suppose that a station A has transmitted a frame, which has not yet reached another station B due to propagation delay. Station B assumes that the channel is idle and transmits its frame. Thus a collision occurs.
- Even if the propagation delay is zero, there will still be collisions. If two stations become ready in the middle of a third station's transmission, both will wait politely until the transmission ends and then both will begin transmitting exactly simultaneously, resulting in a collision.
- **Situation 2:** Suppose that a station A is transmitting while stations B and C are waiting for the transmission to complete. At the instance station A completes transmission, both stations B and C start transmitting simultaneously at the same time. This results is collision.

Advantage of 1-persistent CSMA

- It has better throughput than ALOHA protocols.
- This protocol is far better than pure ALOHA because both stations have the decency to stop from interfering with the third station's frame.
- This approach will lead to a higher performance than pure ALOHA. Exactly the same holds for slotted ALOHA.

Non-persistent CSMA protocol

- A second carrier sense protocol is **nonpersistent CSMA**.
- In this protocol, a conscious attempt is made to be less greedy than in the previous one.

Algorithm

The algorithm of non-persistent CMSA is

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is idle then it transmits the frame immediately.
- If the channel is busy, the station waits for a random time period during which it does not check whether the channel is idle or busy.
- At the end of the waiting time period, it again checks the status of the channel and restarts the algorithm.
- Consequently, this algorithm leads to better channel utilization but longer delays than 1-persistent CSMA.

Advantage of non-persistent CSMA

- Its rate of collision is much less than 1-persistent CSMA. This is because each station waits for a random amount of time before attempting retransmission.
- The probability that multiple stations will wait for same amount of time is extremely low. So, collision between contending stations is greatly reduced.

Disadvantage of non-persistent CSMA

- It reduces the bandwidth usage of network. This is because the channel remains idle even if there are stations who have frames to transmit.
- This occurs since each station wait for a random time before attempting retransmission. There may be multiple stations who are waiting while the channel is idle.

p-persistent CSMA

- The last protocol is **p-persistent CSMA**.
- P-persistent CSMA is an approach of Carrier Sense Multiple Access (CSMA) protocol that combines the advantages of 1-persistent CSMA and non-persistent CSMA.
- In p-persistent CSMA, when a transmitting station has a frame to send and it senses a busy channel, it waits for the end of the transmission, and then transmits with a probability p.
- Since, it sends with a probability p, the name p – persistent CSMA is given.

Algorithm

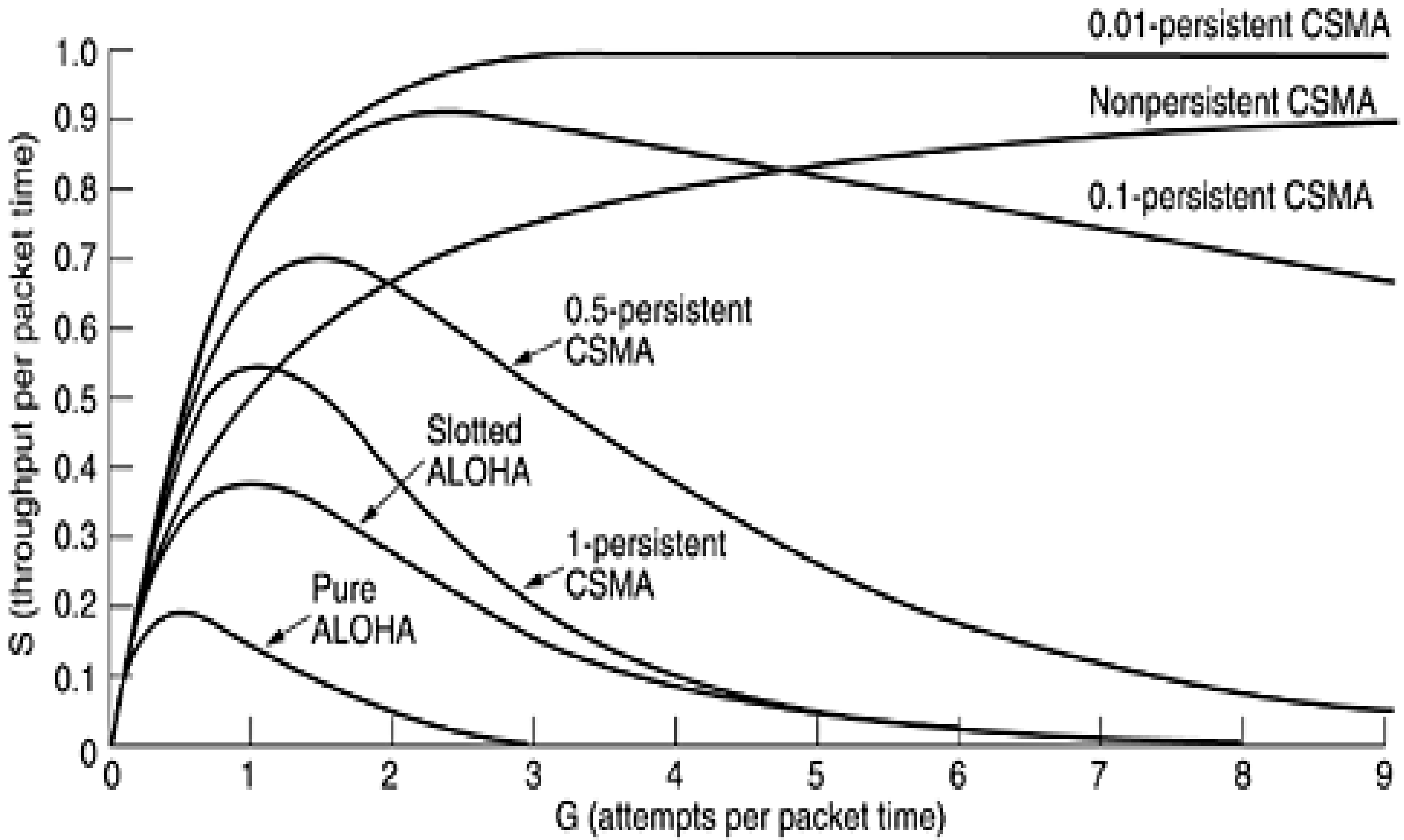
The algorithm of p-persistent CMSA is:

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is idle then it transmits the frame immediately.
- If the channel is busy, the station waits and continually checks until the channel becomes idle.
- When the channel becomes idle, the station transmits the frame with a probability p.
- With a probability (1 – p), the channel waits for next time slot. If the next time slot is idle, it again transmits with a probability p and waits with a probability (1 – p).
- The station repeats this process until either frame has been transmitted or another station has begun transmitting.
- If another station begins transmitting, the station waits for a random amount of time and restarts the algorithm.

Advantage of p-persistent CSMA

- It is the most efficient among 1-persistent CSMA and non-persistent CSMA.
- It reduces the number of collisions considerably as compared to 1-persistent CSMA. The channel utilization is much better than non-persistent CSMA.

Figure 4-4. Comparison of the channel utilization versus load for various random access protocols.



CSMA with Collision Detection

- Persistent and nonpersistent CSMA protocols are clearly an improvement over ALOHA because they ensure that no station begins to transmit when it senses the channel busy.
- Another improvement is for stations to abort their transmissions as soon as they detect a collision.
- In other words, if two stations sense the channel to be idle and begin transmitting simultaneously, they will both detect the collision almost immediately.

CSMA with Collision Detection

- The collision detection technology detects collisions by sensing transmissions from other stations.
- On detection of a collision, the station stops transmitting, sends a jam signal, and then waits for a random time interval before retransmission.
- Quickly terminating damaged frames saves time and bandwidth. This protocol, known as **CSMA/CD** (**CSMA with Collision Detection**) is widely used on LANs in the MAC sublayer.

Algorithms

The algorithm of CSMA/CD is:

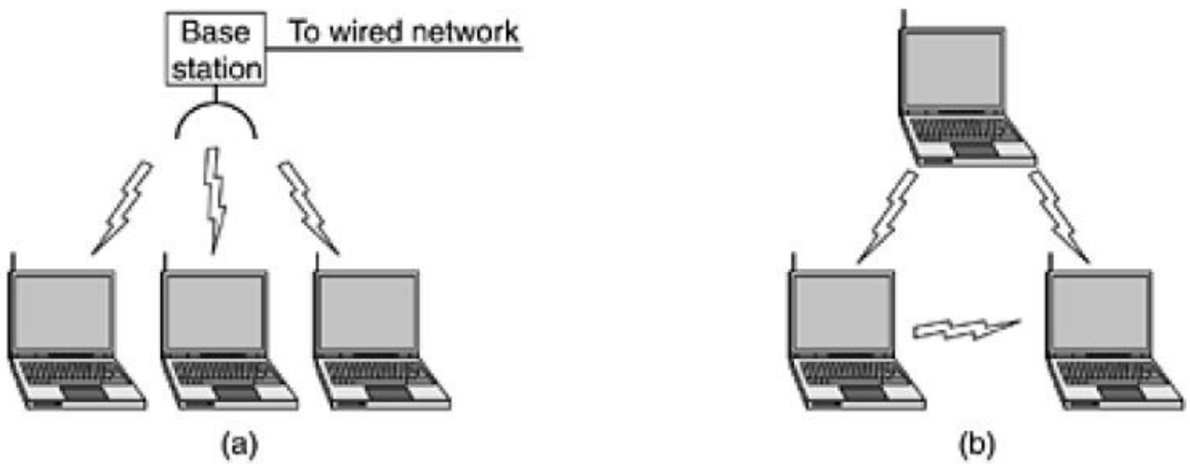
- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is busy, the station waits until the channel becomes idle.
- If the channel is idle, the station starts transmitting and continually monitors the channel to detect collision.
- If a collision is detected, the station starts the collision resolution algorithm.
- The station resets the retransmission counters and completes frame transmission.

The algorithm of Collision Resolution is:

- The station continues transmission of the current frame for a specified time along with a jam signal, to ensure that all the other stations detect collision.
- The station increments the retransmission counter.
- If the maximum number of retransmission attempts is reached, then the station aborts transmission.
- Otherwise, the station waits for a backoff period which is generally a function of the number of collisions and restart main algorithm.

4.2.6 Wireless LAN Protocols

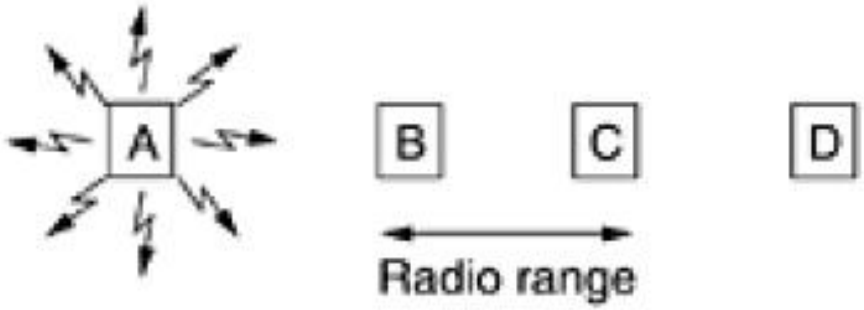
Figure 1-35. (a) Wireless networking with a base station. (b) Ad hoc networking.



- Wireless LANs have somewhat different properties than conventional LANs and require special MAC sublayer protocols.
- A common configuration for a wireless LAN is an office building with base stations (also called access points) strategically placed around the building. All the base stations are wired together using copper or fiber.
- If the transmission power of the base stations and notebooks is adjusted to have a range of 3 or 4 meters, then each room becomes a single cell and the entire building becomes a large cellular system.

- **Assumption :** all radio transmitters have some fixed range. When a receiver is within range of two active transmitters, the resulting signal will generally be garbled and useless.
- It is important to realize that in some wireless LANs, not all stations are within range of one another, which leads to a variety of complications.
- Furthermore, for indoor wireless LANs, the presence of walls between stations can have a major impact on the effective range of each station.
- One approach in wireless LAN might be to try CSMA: just listen for other transmissions and only transmit if no one else is doing so.

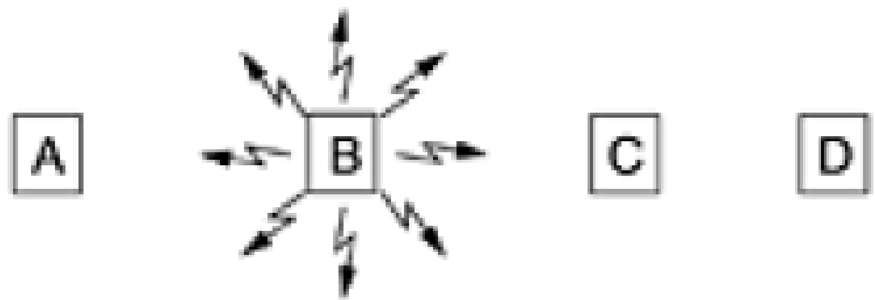
Figure 4-11. A wireless LAN. (a) A transmitting.



(a)

- The radio range is such that station *A and B are within each other's range and can potentially interfere* with one another. *C can also potentially interfere with both B and D, but not with A.*
- **Hidden station problem:** when *A is transmitting to B, If C* senses the medium, it will not hear *A because A is out of range, and thus falsely conclude that* it can transmit to *B. If C does start transmitting, it will interfere at B, wiping out the frame* from *A.*
- *The problem of a station not being able to detect a potential competitor for the* medium because the competitor is too far away is called the **hidden station problem.**

(b)* B *transmitting.



(b)

Exposed station problem: Now let us consider the reverse situation: *B transmitting to A, If C* senses the medium, it will hear an ongoing transmission and falsely conclude that it may not send to *D, when in fact such a transmission would cause bad reception only in the zone* between *B and C, where neither of the intended receivers is located. This is called the* **exposed station problem.**

- The problem is that before starting a transmission, a station really wants to know whether there is activity around the receiver.
- CSMA tells it whether there is activity around the station sensing the carrier.
- With a wire, all signals propagate to all stations so only one transmission can take place at once anywhere in the system.
- In a system based on short range radio waves, multiple transmissions can occur simultaneously if they all have different destinations and these destinations are out of range of one another.

MACA and MACAW

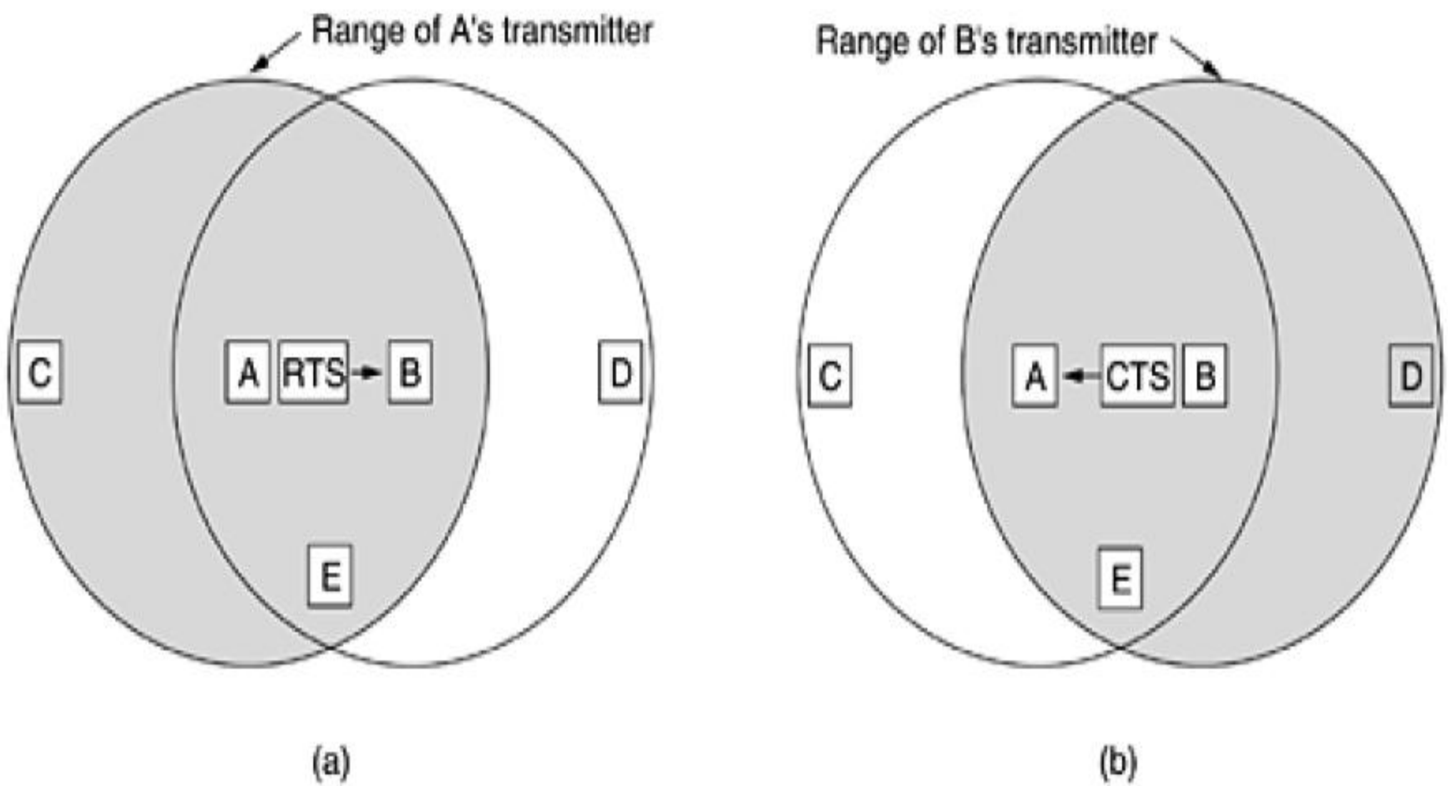
- Protocol designed for wireless LANs is:

❑ **MACA (Multiple Access with Collision Avoidance) Protocol**

❑ **MACAW (MACA for Wireless) Protocol**

MACA protocol

Figure 4-12. The MACA protocol. (a) A sending an RTS to B. (b) B responding with a CTS to A.



- The basic idea behind it is for the sender to stimulate the receiver into outputting a short frame, so stations nearby can detect this transmission and avoid transmitting for the duration of the upcoming (large) data frame.
- Let us now consider how *A sends a frame to B. A starts by sending an* ***RTS (Request To*** **Send)** frame to *B.* This short frame (30 bytes) contains the length of the data frame that will follow.
- Then *B replies with a* ***CTS (Clear to Send)*** *frame.* The CTS frame contains the data length (copied from the RTS frame).
- Upon receipt of the CTS frame, *A begins transmission.*

- Now let us see how stations overhearing either of these frames react. Any station hearing the RTS is clearly close to *A and must remain silent long enough for the CTS to be transmitted* back to *A without conflict.*
- *Any station hearing the CTS is clearly close to B and must remain* silent during the upcoming data transmission, whose length it can tell by examining the CTS frame.

- In Fig., *C is within range of A but not within range of B. Therefore, it hears the RTS from A but not the CTS from B.*
- *As long as it does not interfere with the CTS, it is free to transmit* while the data frame is being sent.
- In contrast, *D is within range of B but not A. It does not* hear the RTS but does hear the CTS. Hearing the CTS tips it off that it is close to a station that is about to receive a frame, so it defers sending anything until that frame is expected to be finished.
- Station *E hears both control messages and, like D, must be silent until the data frame* is complete.

- Despite these precautions, collisions can still occur. For example, *B and C could both send RTS* frames to *A at the same time. These will collide and be lost.*
- *In the event of a collision, an* unsuccessful transmitter (i.e., one that does not hear a CTS within the expected time interval) waits a random amount of time and tries again later.

MACAW (MACA for Wireless)

- Bharghavan et al. (1994) fine tuned MACA to improve its performance and renamed their new protocol **MACAW (MACA for Wireless). It includes:**

❑ **An ACK frame** after each successful data frame.

❑ **Backoff algorithm separately for each data stream (source-destination pair)**, rather than for each station. This change improves the fairness of the protocol.

❑ **A mechanism for stations to exchange information about congestion**

MACAW (MACA for Wireless)

- Bharghavan et al. (1994) fine tuned MACA to improve its performance and renamed their new protocol **MACAW (MACA for Wireless).**
- They noticed that without data link layer acknowledgements, lost frames were not retransmitted until the transport layer noticed their absence.
- They solved this problem by **introducing an ACK frame** after each successful data frame.
- They also observed that CSMA has some use, namely, to keep a station from transmitting an RTS at the same time another nearby station is also doing so to the same destination, so carrier sensing was added.
- In addition, they decided to **run the backoff algorithm separately for each data stream (source-destination pair)**, rather than for each station. This change improves the fairness of the protocol.
- Finally, they **added a mechanism for stations to exchange information about congestion** and a way to make the backoff algorithm react less violently to temporary problems, to improve system performance.