Medium Access Protocols for a Wireless Channel[1]

- In comparison with a fixed channel, a wireless channel is
  - unreliable
  - subject to fading phenomena (slow fading and fast fading)
  - susceptible to interference from other channels and other kind of disturbances
- Bit error rate of a wireless channel is non-negligible
  - this causes special problems for the TPC flow control protocol as this interprets packet losses due to bit errors as a sign of congestion and reduces the window size
  - this issue will be studied elsewhere
- Here we study performance problems that arise when several users attempt to use the same channel.
- A distributed medium access (MAC) protocol is needed to coordinate the actions of different users (nodes).
  - if the number of nodes is large, a *controlled* protocol may not be feasible
  - in such cases simple *random* MAC protocols are used

---

[1] Largely based on: A. Kumar, D. Manjunath, J. Kuri, Communication Networking, Elsevier, 2004; J. Hammond, P. O'Reilly, Local Computer Networks, Addison-Wesley, 1986; L. Kleinrock, Queueing System, Vol. II, Addison Wesley, 1976.

Aloha protocol

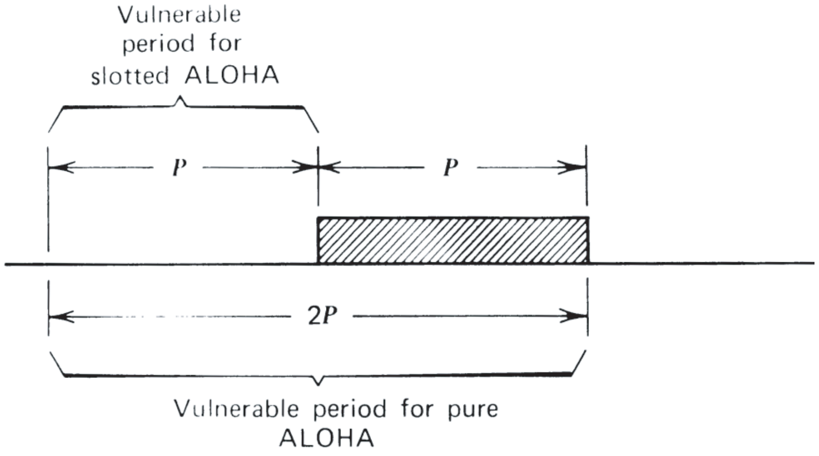
- Random access procedures were first developed for long radio links typical, e.g., in satellite communications
  - later they were adopted for communication in bus-type medium: the Ethernet
  - today they are common in all kinds of wireless data communication systems
- The delay of getting feedback from the channel is long.
- It would be wasteful if the users had to wait to hear about the success of the transmission before being able to transmit another packet.
- Pure Aloha is the simplest protocol one can imagine: If a node has a packet to transmit, it transmits!
- The protocol does not exclude the possibility of collisions
  - overlapping transmission cannot be decoded properly
  - collided packets have to be resent
  - however, not until a random waiting time; otherwise they would collide again

Performance of pure Aloha

- To analyze the throughput performance pure Aloha we make the following assumptions
  - all the packets have the same length
  - for convenience, the transmission time of a packet is selected as the time unit so that the transmission time of all the packets is 1
  - the number of nodes is large and the traffic rate from each node is low
  - the arriving fresh packet stream can then be modelled as a Poisson process
  - resent packets are delayed for a random time long enough so that they are mixed with the fresh packets without causing any noticeable correlation effect
  - thus the total packet stream can be assumed Poissonian
- The number of transmission attempts in unit (transmission) time is Poisson distributed with mean G (including both fresh and retransmitted packets).

Performance of pure Aloha, cont.

- In pure Aloha, the contention or vulnerability period is two units long, see the figure.
- The transmission of a packet is successful if during the contention period no other packet transmission attempts occur. This happens with the probability e^{-2G}.
- Since the number of attempts per time is G the rate of successful transmission per unit time or the throughput S is
$$S = G\,e^{-2G}$$
- One easily finds that the maximum throughput is obtained when $G = \frac{1}{2}$ and yields a very low maximum efficiency $1/2e \approx 0.184$; compare with the ideal maximum of 1 corresponding to successful transmission of packets back-to-back.



From L. Kleinrock, Queueing System, Vol. II, Addison-Wesley, 1976.

Slotted Aloha (S-Aloha)

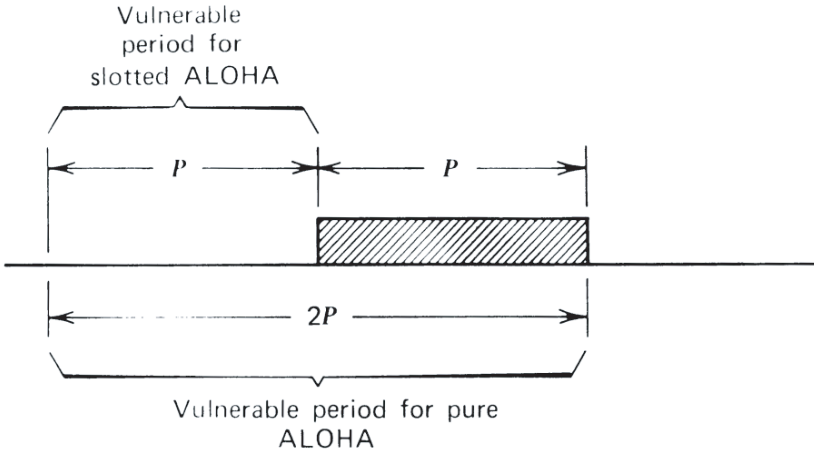
- Slotted Aloha is a refinement of pure Aloha to obtain a better performance.
- Time is segmented into slots of a fixed length, equal to the packet transmission time.
- Transmissions are forced to collide either completely or not at all
  - transmission of a generated packet is delayed so that it will fit exactly in the next slot
  - requires additional overhead to provide the time synchronization information
- The length of the vulnerability period is reduced from 2 to 1, improving the performance.
- Slotted Aloha is used in, e.g.,
  - GSM systems for sending control messages from the mobile terminals to the base station in so-called RACH channel (Random Access Channel)
  - VSAT (Very Small Aperture Terminal) satellite network for sending channel reservation messages
  - data transmission over cable-TV network using the DOCIS (Data Over Cable Service Interface Specifications) standard for sending channel reservation requests from the cable modem to the head end

Slotted Aloha cont.

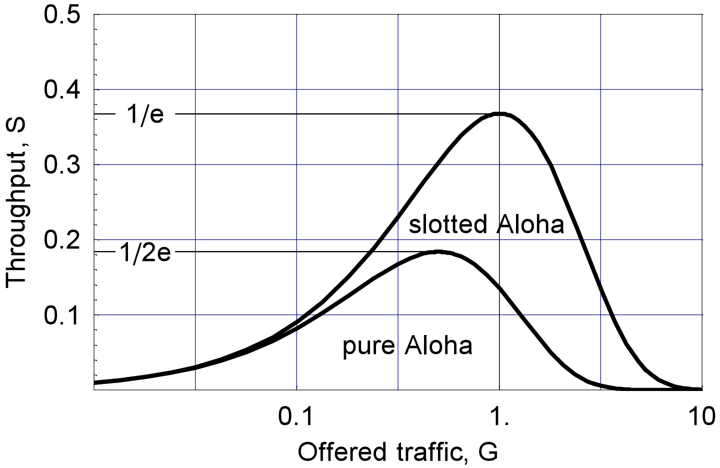
- In slotted Aloha, the vulnerability period is one unit long, see the figure.
- The mean number of packets intended to be sent (transmission attempts) in a slot is G; the actual number is Poisson distributed with this mean.
- A successful transmission occurs when there is exactly one transmission attempt in a slot. Therefore the throughput S is

$$S = G\, e^{-G}$$

- The maximum throughput is obtained when $G = 1$ and is still low $1/e \approx 0.368$.
- Throughput as a function of G is shown in the figure below.



From L. Kleinrock, Queueing System, Vol. II, Addison-Wesley, 1976.



Instability of Aloha protocol

- A peculiar feature of the throughput curve of Aloha protocol is that it has a maximum at a given value of G ($G = \frac{1}{2}$ for pure Aloha and $G = 1$ for slotted Aloha).
- Beyond that value the throughput is a decreasing function of G rendering the system unstable:
  - a small increase in the transmission attempt rate *decreases* the throughput, leading to more backlogged packets, more retransmissions thus further increasing G and decreasing the throughput...
  - ultimately the throughput goes to zero and the backlog grows indefinitely; all slots are wasted by colliding transmissions of the retransmitted packets
- In the following we study this instability in more detail for slotted Aloha and discuss some methods designed to stabilize the algorithm.

Instability of slotted Aloha

- In order to analyze the instability we have to consider the *dynamic* behaviour of the system, i.e., how the queue of backlogged packets behaves in time.
- A packet that has suffered a collision stays in the network and makes retransmission attempts until successful; such packets are called backlogs.
- Assume that
  - fresh arrivals during a slot will always attempt a transmission at the beginning of the next slot
  - backlogs attempt a retransmission in each slot with probability p_r; that is, the time to retransmission attempt is geometrically distributed
- Denote

$$\begin{cases} B_k & = & \text{backlog at the beginning of slot } k \\ A_k & = & \text{number of new arrivals in slot } k\text{; } A_k \sim\text{Poisson}(\lambda) \\ D_k & = & \text{number of departures in slot } k\text{, } D_k \in \{0,1\} \end{cases}$$

- Obviously

$$B_{k+1} = B_k + A_k - D_k,$$

and B_k constitutes a discrete time Markov chain.

Instability of slotted Aloha, cont.

- Consider the *drift* $d(n)$ defined as the expected change in the backlog given that the current backlog is n,

$$d(n) = \mathrm{E}\left[B_{k+1} - B_k \,|\, B_k = n\right] = \mathrm{E}\left[A_k - D_k \,|\, B_k = n\right].$$

- Given that the current backlog is n, the backlog
  - decreases by 1, if no new arrival occurs and if only one of the backlogs attempts transmission
  - increases by one if exactly one arrival occurs and if at least one of the backlogs attempts transmission
  - increases by amount $m \geq 2$ if the number of new arrivals is m
  - remains unchanged in all other cases

Instability of slotted Aloha, cont.

- Thus we can write

$$\begin{cases} \mathrm{P}\{B_{k+1}-B_k=-1\,|\,B_k=n\} &=& e^{-\lambda}\,n\,p_r\,(1-p_r)^{n-1} \\ \mathrm{P}\{B_{k+1}-B_k=+1\,|\,B_k=n\} &=& \lambda\,e^{-\lambda}\,(1-(1-p_r)^n) \\ \mathrm{P}\{B_{k+1}-B_k=m\,|\,B_k=n\} &=& \dfrac{\lambda^m}{m!}\,e^{-\lambda}, \quad \text{for } m \geq 2 \end{cases}$$

- Using these and simplifying we get

$$\begin{aligned} d(n) &= \mathrm{E}\,[B_{k+1}-B_k\,|\,B_k=n] = \sum_{m=-1}^{\infty} m\,\mathrm{P}\{B_{k+1}-B_k=m\,|\,B_k=n\} \\ &= (-1)\,e^{-\lambda}\,n\,p_r\,(1-p_r)^{n-1}+(+1)\,\lambda\,e^{-\lambda}\,(1-(1-p_r)^n)+\sum_{m=2}^{\infty} m\,\frac{\lambda^m}{m!}\,e^{-\lambda} \\ &= \lambda-e^{-\lambda}\,(1-p_r)^n\left(\lambda+\frac{n\,p_r}{1-p_r}\right) \end{aligned}$$

- Because of the factor $(1-p_r)^n$, the second term becomes negligible for large n; thus the mean drift for all large values of n (beyond a finite threshold) is positive. For large n the network has a tendency to increase the backlog rather than decrease it. It can develop a large backlog that may never be cleared. Thus Aloha protocol with fixed retransmission probability p_r is unstable.

Stabilizing the Aloha algorithm

- Several protocols have been devised that stabilize the Aloha system.
- The idea in these protocols is to make the retransmission probability adaptive, decreasing it when the backlog (either directly observed or indirectly inferred) grows.
- We first consider one such strategy, a rather theoretical one, and then discuss a more practical approach.

Ideal retransmission policy

- A successful transmission occurs either a) if one packet arrives and none of the n backlogs attempt retransmission or b) if no new packet arrives and exactly one of the n backlogs attempts retransmission.
- The probability of success, P_s, is thus

$$P_s = \lambda\, e^{-\lambda}\, (1 - p_r)^n + e^{-\lambda}\, n\, p_r\, (1 - p_r)^{n-1}$$

- The optimal, state-dependent retransmission probability $p_r(n)$ that maximizes this is

$$p_r(n) = \frac{1-\lambda}{n-\lambda}$$

- So, if we assume (unrealistically) that each node knows the size of backlogs n and the arrival rate λ, this is the best choice for the retransmission probability.
- If this adaptive retransmission probability is used, the drift becomes

$$d(n) = \lambda - e^{-\lambda} \left(\frac{n-1}{n-\lambda}\right)^{n-1}$$

- It is easily seen (exercise) that $d(n) \to \lambda - 1/e$ when $n \to \infty$. This means that for arrival rates $\lambda < 1/e$ the system is stable.

Realistic retransmission policy – back-off mechanism

- It is not practical for the nodes to know the size of the backlog.
- In many random access protocols, nodes use their transmission history to adapt retransmission protocols.
- In particular, an unsuccessful transmission attempt leads to a decrease in the retransmission probability – so-called back-off.
- A typical back-off algorithm is as follows:
  - after a collision is detected, a new transmission is attempted only after a back-off period x
  - the length x of the back-off period is drawn uniformly in the interval $(0, B-1)$
  - B is updated at every transmission attempt:

$$B = \begin{cases} \min\left(a \times B, B_{\max}\right) & \text{if transmission collides} \\ \max\left(B - b, B_{\min}\right) & \text{if transmission is successful} \end{cases}$$

    where $a,\ b,\ B_{\min}, B_{max}$ are predefined constants

  - often $a = 2$, whence this is called binary exponential back-off

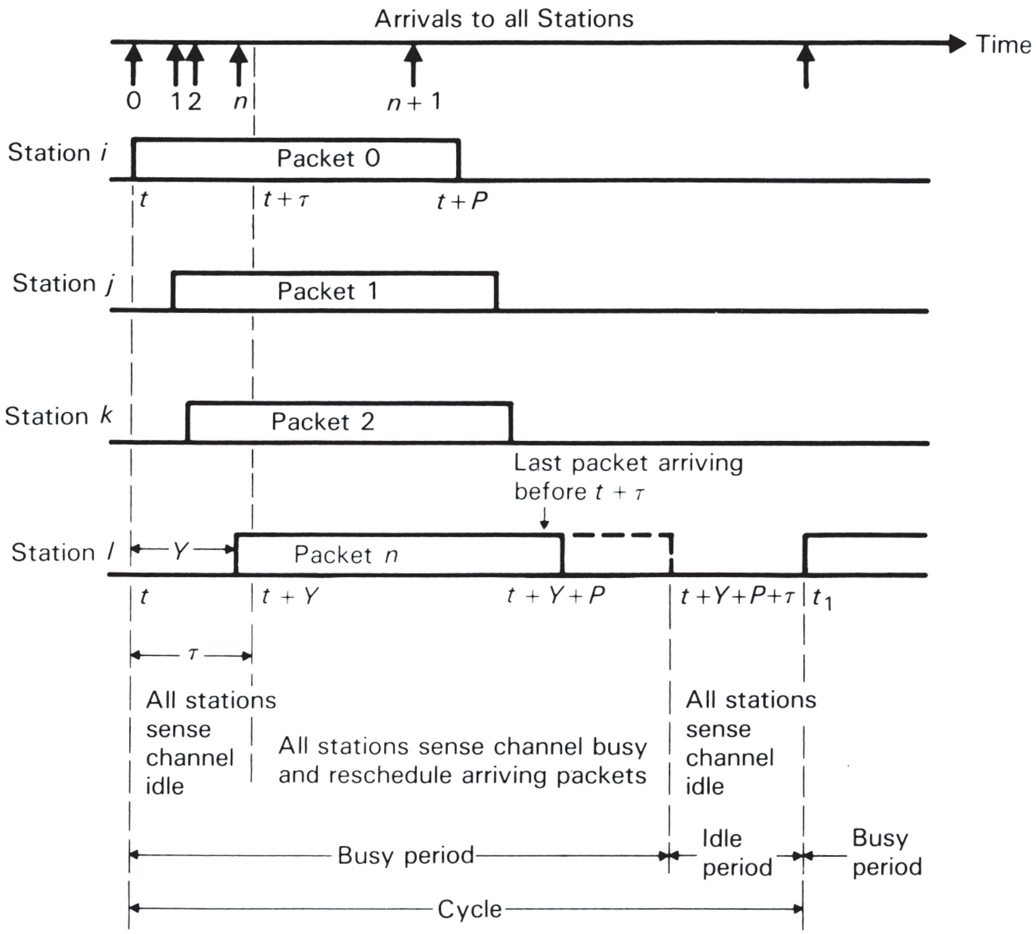
Carrier Sense Protocols, CSMA (Carrier Sense Multiple Access)

- Here we consider networks having small propagation delays compared with the packet transmission time
    - compact terrestrial radio network (as opposed to, e.g., satellite communication)
    - wired bus type network, like the Ethernet
- It is feasible for a station to “listen” to the channel to determine whether it is busy before a transmission is attempted and defer the transmission until the channel is sensed idle.
    - additional hardware is needed for this, though
- CSMA-type protocols are more efficient than Aloha or S-Aloha.
- Basically CSMA protocol works in continuous time
    - or, if slotted, the slot granularity is much finer than the packet transmission time, so that in effect it looks almost like a continuous time system
- The efficiency of the CSMA protocol stems from the facts that
    - the vulnerability (collision) period is much shorter than with Aloha or S-Aloha
    - the packet does not have to wait for the next slot boundary in order to be transmitted as in S-Aloha; less time is wasted

Carrier Sensing Protocols cont.

- As with the Aloha, there is a vulnerability period during which a collision may occur.
- The length of the vulnerability period is equal to τ, the one-way propagation delay
  - as shown in the figure, when a station starts to transmit, because of the propagation delay, in the worst case it may take time τ before other stations get informed about the on-going transmission; all packets whose transmission starts during this period collide with the initial packet
- In CSMA, once the transmission is started, the whole packet is sent
  - the sending station does not "know" if other packets collide
  - the contrary is true in systems with collision detection (CD), as discussed later
- After the last of the collided packets has ended, it takes again time τ until all stations sense the channel idle and can attempt transmitting again
- To summarize:
  - collision occurs if during the initial period τ at least one transmission is attempted
  - once a collision occurs, the system is useless for a total time of $P + 2\tau$, where P is the packet transmission time

Carrier Sensing Protocols cont.



Components of a cycle containing an unsuccessful busy period for nonpersistent CSMA.

From J. Hammond and P. O'Reilly, Local Computer Networks, Addison-Wesley, 1986.

Variations of the CSMA Protocol

- When a packet collides (ultimately learned by a missing ack), the transmission is always rescheduled to a later time using some specified back-off algorithm.
  - after the back-off, the station again senses the channel and repeats the algorithm
- At some point, the station has a packet ready to transmit
  - the station is called ready, irrespective whether the packet is a new or a retransmission
- There are some variations of the CSMA protocol depending what a ready station does in finding the channel busy/idle
  - nonpersistent CSMA
  - p-persistent CSMA

Nonpersistent and p-persistent CSMA Protocols

- The *nonpersistent* CSMA works as follows
  - if the channel is sensed idle, the packet is transmitted
  - if the channel is sensed busy, the node waits a random amount of time (back-off), senses the channel again and the algorithm is repeated
- The *p-persistent* CSMA works as follows
  - if the channel is sensed idle, then with probability p the node transmits the packet; with the probability $(1 - p)$, the node waits time τ (propagation delay), and the algorithm is repeated
  - if the channel is sensed busy, the node persists in sensing the channel until it becomes idle operates as in the previous step (channel sensed idle)
- A special case of the *p-persistent* CSMA protocol is 1-persistent CSMA, where a ready station always begins transmission when sensing the channel idle; persistently senses a busy channel and starts transmitting immediately when sensing the channel idle
  - the Ethernet and the IEEE 802.3 LAN and IEEE 802.11 WLAN MAC protocols are 1-persistent

Performance of the CSMA Protocols

- The performance analysis of the CSMA protocols is straight-forward but somewhat tedious and not very illustrative.
- Here we only give some results without derivation.
- *Throughput of nonpersistent CSMA*

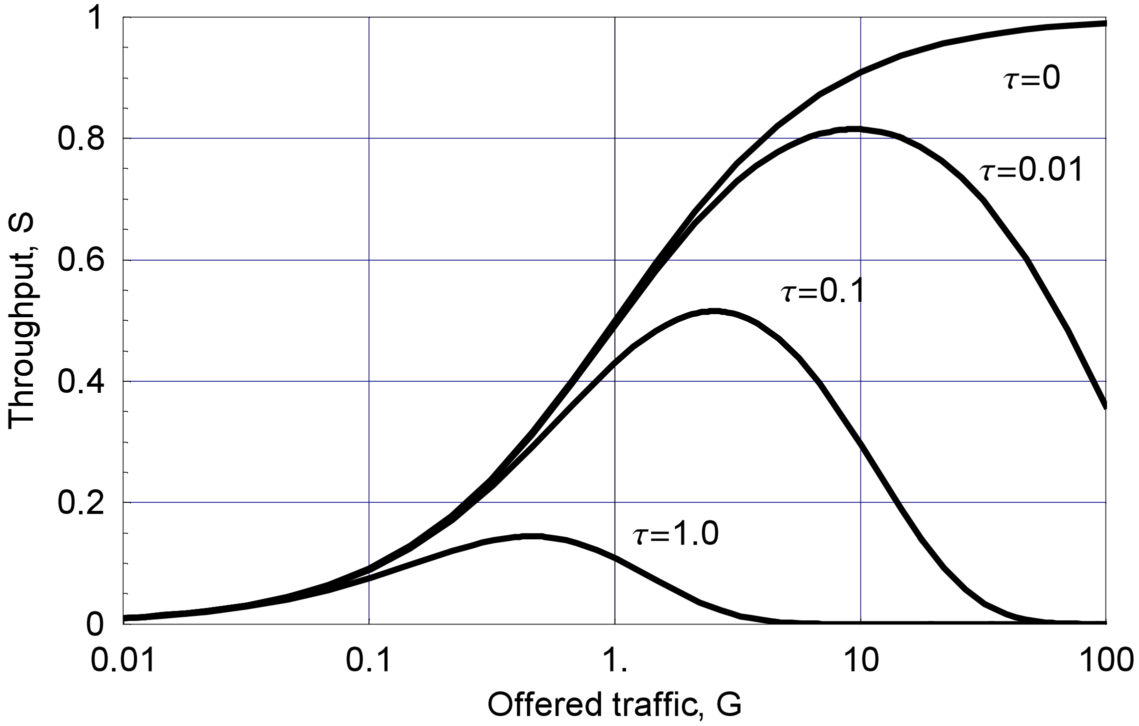
$$S = \frac{Ge^{-\tau G}}{G(1+2\tau) + e^{-\tau G}}$$

- *Throughput of 1-persistent CSMA*

$$S = \frac{G[1 + G + \tau G(1 + G + \tau G/2)]e^{-(1+2\tau)G}}{G(1+2\tau) - (1 - e^{-\tau G}) + (1 + \tau G)e^{-(1+\tau)G}}$$

- Note that, as before, time is measured in terms of the packet transmission time,
  - τ is the ratio of one-way propagation delay to the packet transmission time
  - this parameter is assumed small when CSMA is used
  - G is the traffic load, average number of arrivals in the transmission (service) time

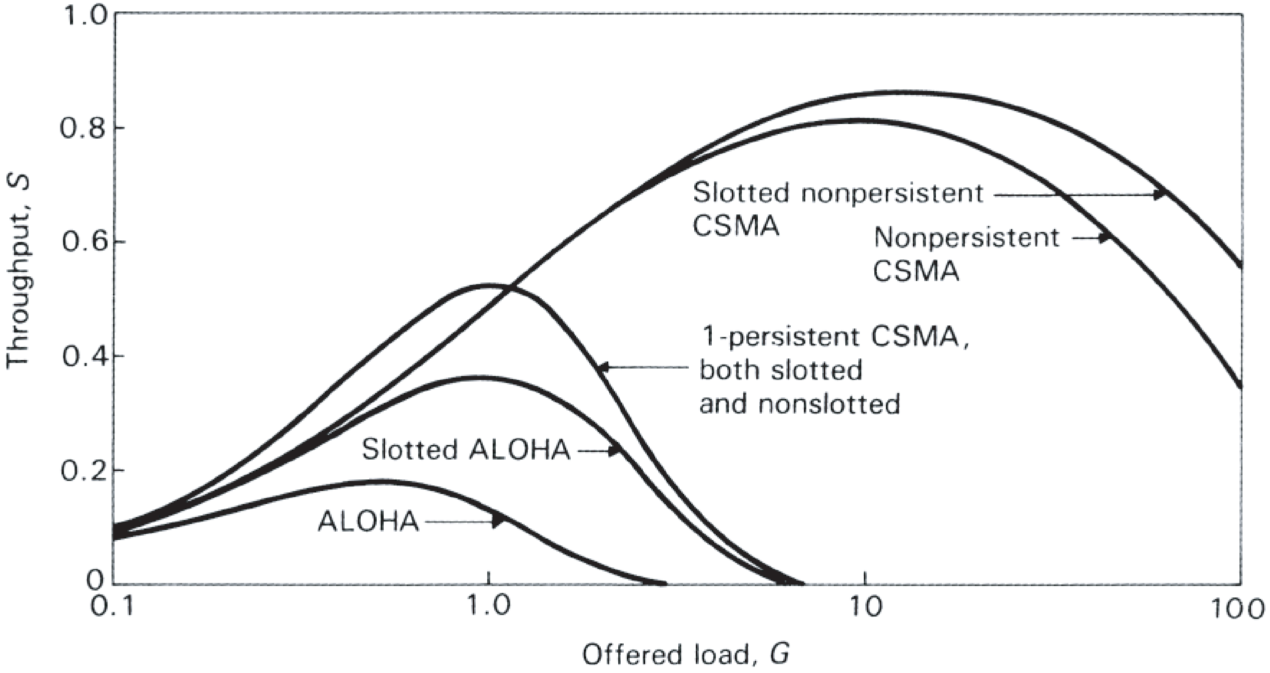
Throughput of nonpersistent CSMA for different values of τ



Throughput S versus load G.

Comparison of the throughput performance of different CSMA protocols

- In these comparisons value $\tau = 0.01$ is assumed for the one-way propagation delay.



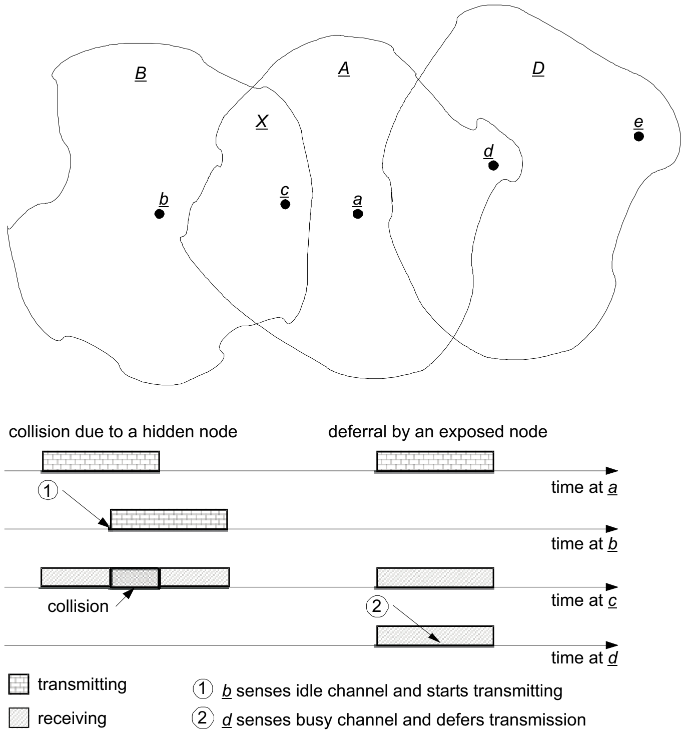
From J. Hammond and P. O'Reilly, Local Computer Networks, Addison-Wesley, 1986.

CSMA with Collision Detection, CSMA/CD

- To further improve CSMA, the node can continue to monitor the channel after beginning transmission to detect a possible collision
  - again needs additional hardware
  - is technically more demanding than just sensing that the channel is busy (may be infeasible in wireless networks)
- If collision is detected, the transmission is immediately stopped to minimize the waste of the channel capacity.
- CSMA/CD was invented for the popular Ethernet local area network (IEEE 802.3).
- The Ethernet uses the back-off scheme given before with the parameters $a = 2$ (binary exponential back-off), $b = B$, $B_{\min} = 2$, $B_{\max} = 1024$
  - the unit of back-off period in Ethernet is equal to twice the maximum round-tip delay in the network (in 10 and 100 Mbps Ethernet equal to 512-bit transmission times).

Hidden and exposed terminal problems

- Spatial reuse of frequency spectrum in wireless networks
  - nodes in different parts of the network can send simultaneously (in the same band)
- So-called hidden and exposed terminals cause problems for the MAC protocol.
- *Hidden terminal* refers to the case where a receiving node c is within the transmission range of two nodes a and b but these are unable to sense each other's transmissions
  - they may transmit simultaneously to c unaware of collision occuring at c
  - a is hidden from b and vice versa
- *Exposed terminal* refers to the case where node d wants to transmit to node e ($d \rightarrow e$) but an ongoing transmission $a \rightarrow c$ is sensed by d. Node d unnecessarily defers its transmission even though reception at e would not be interfered by a. Node d is exposed to node a.



The transmission ranges of nodes a, b and d are shown as (irregular, to be realistic) regions A, B and D.

From A. Kumar, D. Manjunath, J. Kuri, Communication Networking, Elsevier, 2004.

Collision avoidance by the RTS/CTS handshake

- A possible solution to the hidden terminal problem is the use of a *busy tone*
  - a narrowband auxiliary signalling channel is defined
  - a node actively receiving data transmits busy tone to let the potential other sending nodes know about the ongoing transmission
- Dividing the available spectrum in two parts may be cumbersome. Therefore, in WLANs a different strategy is adopted.
- Actual data transfer is preceded by a *handshake* between the transmitter and the receiver
  - a node wanting to send data first transmits a short *request to send* (RTS) packet
  - if the destination receives the RTS correctly and is free to accept data, it acknowledges the request by a *clear to send* (CLS) packet
  - if the CLS is not received within a specified timeout period, a retransmission is attempted after a random back-off period
  - after a successful RTS/CTS exchange, the channel is reserved for data transfer
- This called *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA).

Discussion of the CDMA/CA scheme

- Packet length information (of data) is included in the RTS/CTS packets
  - other nodes can determine the time to completion and schedule their transmissions accordingly
  - they must not try to send their RTS packets during the data transfer
- To completely eliminate the hidden terminal problem, the transmission range of the CTS packet should be larger than the interference range of other nodes
  - this may not be true in practice; some interfering nodes may not hear the CTS at all or may hear it but not able to decode it
  - such nodes may transmit their RTS packets during the data transfer causing a collision
- The RTS/CTS scheme does not at all address the exposed terminal problem
  - the problem is indeed difficult to solve
  - even if the exposed node would be allowed to send its RTS, it could not itself receive the subsequent CTS
  - does not prevent the wireless network operating but causes performance degradation

Multiple Access in IEEE 802.11

- The 802.11 specifications define two modes of operation
  - *Point Coordination Function* (PCF), centralized polling-based
  - *Distributed Coordination Function* (DCF), distributed MAC
- The DCF random access procedures are based on the CSMA/CA mechanism with RTS/CTS packets
- The figure shows the events during data transfer between four nodes.

From A. Kumar, D. Manjunath, J. Kuri, Communication Networking, Elsevier, 2004.

- At the end of data transfer, there is a short interframe space (SIFS) to allow the receiving node turn around its radio and send an ACK-packet.
- When channel is sensed idle, before sending RTS every node waits a time, DCF interframe space (DIFS > SIFS), to allow the ACK to capture the channel.