2 MULTIPLE ACCESS TECHNIQUES

In any communication system one of the main problems is the way in which the ensemble of users have access to the communication channel in order to transmit and receive information in both directions without interfering one another.

Mobile communications systems, are, by definition, systems with more then one mobile users that transmits and receive data in both directions**.** The communication between those mobile terminals is made through a base station, which receives the data from all users within its covering area, transmits it to the destination and controls the way in which those information are transmitted, so the system functionality is ensured.

There are two issues that has to be solved in mobile communications systems: the simultaneous exchange of data between the mobile terminal and the base station and the method used to allow multiple users to share the system.

2.1. DUPLEX TRANSMISSIONS

As it has been mentioned above, one of the main requirements imposed to the system is to ensure that the mobile terminal can simultaneously transmit and receive data to and from the base station. This bidirectional communication is named *d**uplex transmission.*** As terminology,

- The link between the base and the mobile station is called ***downlink*** in European terminology and ***forward link*** in U.S. terminology;
- The link between the mobile station and the base station is called ***uplink*** in European terminology and ***reverse link*** in U.S. terminology;

The duplex transmission can be achieved using frequency or time division, as suggested in figure 2.1.

- the ***frequency division duplexing*** (FDD) technique offers two frequency bands for each user, one for uplink and one for downlink. Each duplex channel consists of two simplex channels, each of them being transmitted on a different frequency carrier and using another frequency band. Both mobile and base stations have to be equipped with two antennas (one for each link) and a duplexer (consisting of radio frequency filters which protect the receiver from interference with adjacent channels) which permits sending and receiving messages simultaneously through both channels. For most mobile radio systems the frequency separation between the uplink and the downlink channels is a constant of the system and does not depend on the type of channel used..

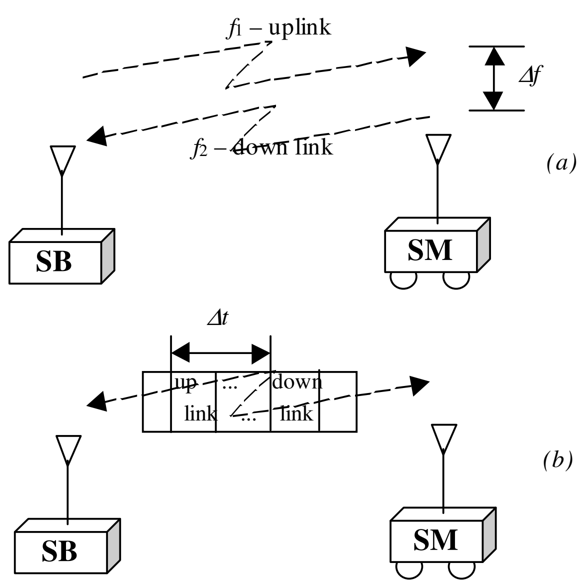

Fig. 2.1. The description of FDD (a) and TDD (b) techniques

- the ***time division duplexing*** (TDD) technique use the same frequency band for both links. The data are organized packets and use different time slots for uplink and downlink. If the time difference between the two allocated slots is relatively small, the transmission and reception of data appears to be simultaneous from the users point of view. This type of duplexing is frequently used in digital communications systems, in which data can easily organized in packets.

The FDD duplexing technique is frequently used in analog radio communications, in which the resources offers enough spectrum to allot different carriers and bandwidths for the two links. A power control scheme has to be implemented in order to avoid interchannel interference. Moreover, the frequencies and bandwidths has to be chosen in a proper manner, to reduce the costs of the receiver (including the duplexers, transceivers and antennas for both mobile and base stations).

The TDD duplexing techniques allows the mobile and base stations to operate as both transmitter and receiver, using the same carrier frequency and bandwidth. The separations between uplink and downlink is made in time domain. Hence, the transmission cannot be simultaneous in both directions, but it appears as “quasi simultaneous” if the time difference between the slots is small. The main issue in TDD technique is the time synchronization between transmitter and receiver; if this synchronization is not perfect, inter-channel interference occurs.

2.2. MULTIPLE ACCESS IN MOBILE COMMUNICATIONS

In mobile communication systems, both analogue and digital, multiple users must have access to the system resources in order to transmit and receive data with one or several base stations. In order to achieve that, three fundamental multiple access techniques are usually used:

- frequency division multiple access – **FDMA**;
- time division multiple access – **TDMA**;
- code division multiple access – **CDMA**.

Mobile communication systems can have a large or narrow bandwidth, depending on the ratio between the bandwidth occupied one single user and the coherence bandwidth of the transmission channel (which represents the frequency range in which the amplitude of two cosine-function components transmitted on different carrier frequencies are correlated one another at the receiver).

2.2.1. Narrowband communication systems

A ***narrowband*** communication system is that system in which the bandwidth used by a channel (user) is smaller or approximately equal to the coherence bandwidth of the transmission channel. In this case the radio frequency spectrum is divided into a large number of narrowband channels, each user being allotted to use such a channel. The multiple access technique can be FDMA or TDMA.

In narrowband FDMA systems, each user has assigned a carrier frequency and a bandwidth which he can use individually the whole transmission time. The duplex communication is achieved by frequency division (FDD). The principle FDMA / FDD is illustrated in figure 2.2.

In narrowband TDMA system, all users access a single communication channel, each of them being assigned a temporal slot in a cyclic mode. Duplexing can be realized in either TDD or FDD technique. The TDMA / FDD systems use two different frequencies, one for the uplink and the other for the downlink. The TDMA / TDD systems use the same carrier frequency and the same frequency bandwidth for both links, the transmission being made using different time slots for the two directions. For this reason the TDD transmission is quasi-simultaneous, because when the data transmission is made in one direction the transmission in the other direction is blocked. If the transmitted data rates are large enough, the time one direction is blocked is negligible and almost imperceptible for classical voice communications. The principle of TDMA is presented in figure 2.3.

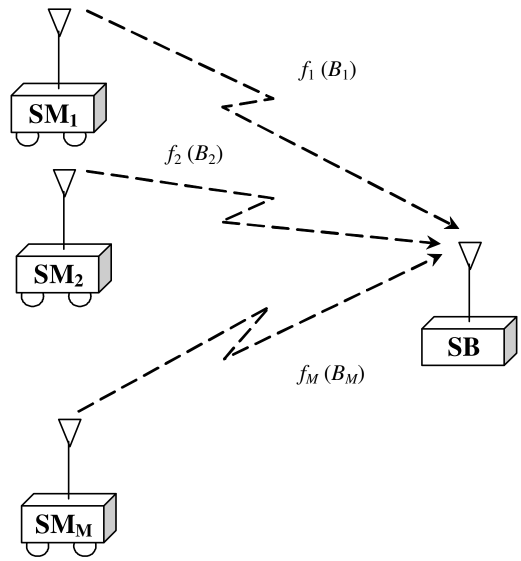

Fig. 2.2a. The principle FDMA / FDD

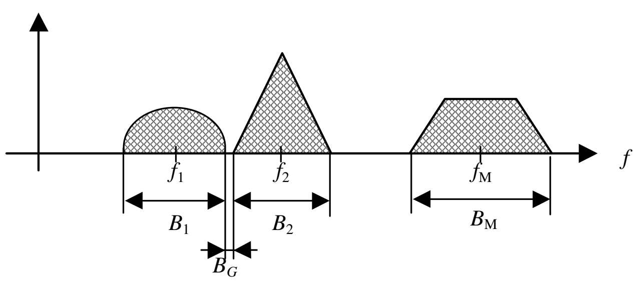

Fig. 2.2b. The frequency spectrum

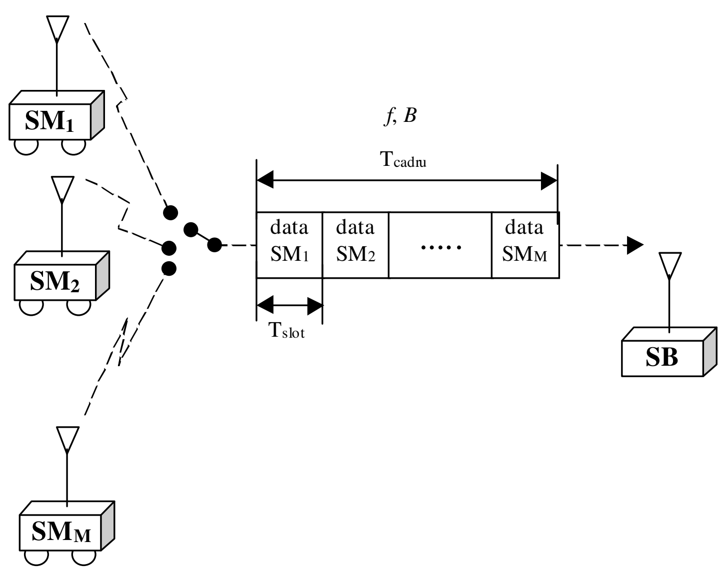

Fig. 2.3. The TDMA principle

Comparison between TDMA / FDD and TDMA / TDD:

- from the spectrum's point of view, the overall bandwidth occupied by the two systems is the same: while the FDD uses two separate frequency bands, the TDD system uses one double bandwidth;
- because the bandwidth used by TDD systems is double compared with the one used by FDD ones, the radio-frequency filters used by the transmitter and receiver have larger bandwidth, so are easier to implement;
- in TDD systems the duplexing is realized using a radio frequency switch that connects the transmitter of receiver antenna, depending on the direction of the data flow. This structure is less complicated then the ones used in FFD duplexing (employing two different antennas and a duplexer).

2.2.2. Broadband communication systems

The ***broadband*** communication systems are those systems in which the bandwidth used by one channel (user) is much larger than the coherence bandwidth of the communication channel. This way the fading due to the multiple propagation paths does not affect as severe the received signal, and the frequency selective fading will affect only a part of spectral components of the transmitted signal.

The multiple access technique used in broadband communication systems is the Code Division Multiple Access (CDMA). In CDMA all users access the *same communication channel* in the *same time*; each used is assigned an individual "spreading" code, orthogonal with all the other codes used by the system, with the rate significantly larger then the data rate. In this way the spectrum of the transmitted signal is spread on a bandwidth proportionally with the code (chip) rate. The bandwidth occupied by the communication channel is several orders higher than the data signal, so the signals are named *spread spectrum signals*.

Depending on the way in which the code combines with the data there are three types of spread spectrum CDMA access systems:

- Direct Sequence – DS-CDMA: the code modulation is performed by multiplying the transmitted data with the spreading code; this technique is largely used in 3rd generation mobile communication systems, due to its technological simplicity;
- Frequency Hopping – FH-CDMA: the carrier frequency is chosen in accordance with the spreading code; this technique is used especially in military communications because is more resistive to jamming and interference, but more expensive because the equipment used is more complex;
- Time Hopping TH-CDMA – uses very short duration Gaussian impulses whose position inside a frame is dictated by the spreading code an the modulation data. This technique, combined with the Pulse Position Modulation (PPM) leaded to Ultra Wide Band (UWB) technique, which is now largely used in short range indoor communication networks.

Beside the ones mentioned above, in mobile communication systems we have to mention the Packet Radio Access technique (PR) derived from the TDMA technique and the Space Division Multiple Access (SDMA) based on the capacity of spatially separate the propagation channel. Those techniques will be analyzed separately, with respect to their capacity and performances.

In table 2.1. are shown several mobile communication standards and the multiple access, respectively the duplexing technique used.

Table 2.1.

Cellular system	Multiple Access / duplexing technique
AMPS (Advanced Mobile Phone System)	FDMA/FDD
GSM (Global System for Mobile)	TDMA/FDD
USDC (US Digital Cellular)	TDMA/FDD
JDC (Japanese Digital Cellular)	TDMA/FDD
CT2 (Cordless Telephone)	FDMA/TDD
DECT (Digital European Cordless Telephone)	FDMA/TDD

IS95 (US Narrowband Spread Spectrum)	CDMA/FDD
UMTS (Universal Mobile Telecommunication System)	CDMA/FDD or CDMA/TDD

2.3. CHARACTERISTICS AND PARAMETERS OF FDMA SYSTEMS

In **FDMA** technique every channel is individualized by a carrier and a frequency bandwidth; this channel is used entirely by the assigned user during its conversation, and cannot be used by any other user. In a time-frequency-code coordination system, the FDMA allocation technique is shown in figure 2.4. The users are separated one another in frequency domain.

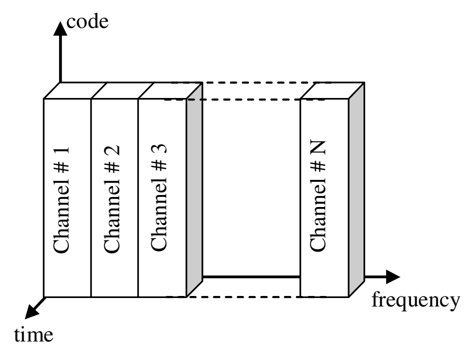

Fig. 2.4. Channel allocation in FDMA systems

The main characteristics of the FDMA systems are :

- each FDMA channel is used by *only one user at a time* instance;
- if the user to whom the channel has been assigned takes a break during the conversation, the channel *is not used* but it cannot be assigned to another user to allow the increase of the overall system capacity ;
- after a certain channel has been allocated to an user, both the base and mobile stations can transmit messages *simultaneously and continuously* ;
- the bandwidth of a FDMA channel is relatively narrow, being equal to the necessary bandwidth of an individual user (usually 30 kHz); the FDMA systems are usually implemented as *narrowband systems* ;

- since the *transmission rate is relatively low*, the symbol period is greater than the propagation delays caused by the channel; therefore the intersymbol interference is relatively low and it does not impose a sophisticated equalization system;
- the *complexity* of FDMA systems *is generally smaller* than the complexity of TDMA systems, regarding the processing necessary for the sent signal ;
- because the FDMA technique allows continuous transmission, it not necessary to transmit a large number of bits for the signalings of the system (for example synchronization or frame splitting) like in the case of TDMA;
- the *costs* necessary for the implementation of FDMA systems are *higher* than in the case of the TDMA ones, because the use of one channel by a single user at a time instance does not lead to a judicious use of resources; moreover, it is necessary the use of some complicated and expensive pass-band filters to reduce the power transmitted outside the assigned bandwidth;
- the FDMA systems require at the receiver pass-band filters with very steep slopes to eliminate cross-talk between channels; those filters make the receiver implementations relatively expensive;
- both base and mobile stations must use duplexer circuits, since the receiver and the transmitter operate in the same time. This leads also to an increase of the cost for the equipment.

The number of channels that can be used simultaneously in a FDMA system is:

$$N = \frac{B_t - 2B_g}{B_c} \tag{2.1}$$

where: - B_t is the frequency band assigned to the system ;

- B_g is the “guard” band needed at the ends of the allocated frequency range;
- B_c is the band of an individual channel.

2.4. CHARACTERISTICS AND PARAMETERS OF TDMA SYSTEMS

In TDMA technique all users use the same band of frequencies, each of them being individualized by a time slot in which they can send or receive messages. The slots are assigned in a cyclic way. In a time-frequency-code coordination system, the FDMA allocation technique is shown in figure 2.5. The users are separated one another in frequency domain.

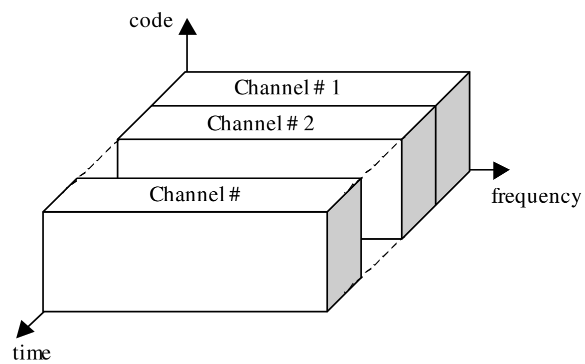

Fig. 2.5. Channel allocation in TDMA systems

The data transmitted from a particular user are kept in a buffer and then send with a rate multiplied by N on the duration of the time slot assigned to that particular user. In this way, from the user point of view, the transmission is discontinuous .

The data transmitted from different users are interleaved into a frame structure. Each frame is consisting of

- a preamble(which contains information for addressing and synchronization),
- a part with data that contains the actual informational messages;
- a number of tail bits used for error detection and correction and for statistical information concerning the quality of the link.
- The data part is also made of slots coming from different users; each slot is also a block of synchronization bits, the data from the message toward one or more users, the tail bits and a number of “guard” bits needed to adjust

the receiver synchronization when crossing from one slot to another or from one frame to another.

A typical TDMA frame structure is shown in figure 2.6.

The **main characteristics** of the TDMA systems are:

- the TDMA systems use *a single carrier frequency* and a *single frequency band*, a temporal slot being assigned for each user.
- the number of slots in each frame depends on the technology, the modulation type used, the assigned frequency bandwidth etc. Transmission of data in the TDMA systems is *not continuous*, but in packets; this leads to a *reduced consume of the mobile station battery*, since a certain user transmits data only on the duration of the assigned slot.

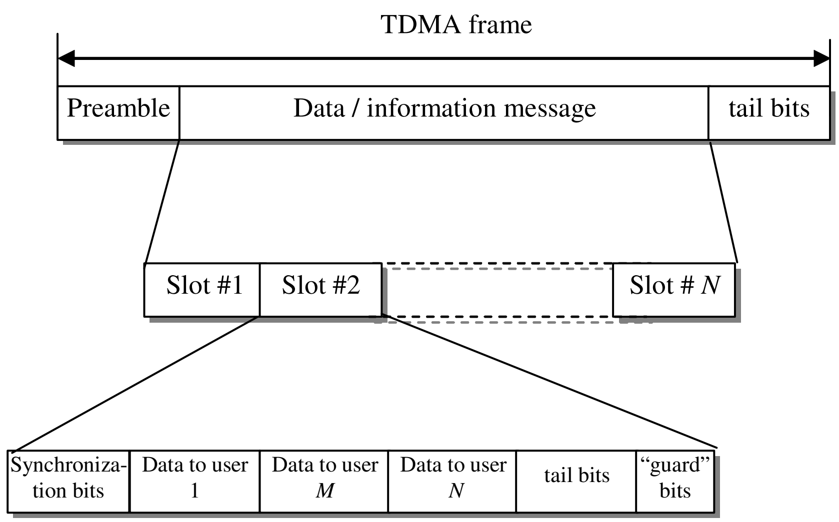

Fig. 2.6. The structure of a TDMA frame

- since the data transmission is made discontinuously, the conversation transfer from one base station to another (*handover)* is performed easier because when the mobile station is inactive it can effectuate the necessary measurements to determine the closest base station; TDMA also allows signal power and error probability monitoring at the reception of each frame

- because an user is using different time slots for transmission and reception it is not necessary to use duplexers in the TDD mode; in FDD mode one can use a simple switch between emitter and receiver; hence the receiver implementation *costs are lower* then in he TDMA case
- since in TDMA the transmission rate *on the common channel* is very high, an *adaptive equalization system* is necessary at the reception in order to reduce the effects of the communication channel, like intersymbol interference;
- because in TDMA the transmission of data is made in packets, is necessary to transmit also a large number of *synchronization bits*, in order to help the receiver to be able to synchronize at the arriving of each packet of data; moreover, a number of *guard bits* are introduced to avoid the superposition of slots that are coming from different users; because of this the supplementary information that must be transmitted beside the useful data, the rate on the common channel is larger then the sum of the rates of all individual users.
- in TDMA is possible to assign *different number of slots* for different users, so the quantity of transmitted data can be modified with respect to the user requirements; in this sense the TDMA systems are more flexible then the FDMA ones;
- TDMA allows the use of a *large domain of transmission rates*, in general multiples of the multiplexer rate (the rate with which the commutation from one user to another is made); therefore, a large variation of procedures and techniques of coding with different bit rate and with different qualities can be used; in this way the price can be chosen by the user in function of the quality imposed by the application.
- the TDMA systems *can be completely realized in digital technology* by large scale integration (VLSI) without using radio frequency filters of small bandwidth which leads to substantial smaller cost

Te ***efficiency of a TDMA frame*** is defined as a measure of the data percentage effectively containing information from the total data that is transmitted in a frame. For this we must first determine the number of supplementary bits transmitted in each frame.

$$b_{\sup} = N_r\left(b_r + b_g\right) + N_t\left(b_p + b_g\right) \tag{2.2}$$

where :

- N_r is the number of reference packets (i.e. internal signaling of the system) by frame;
- N_t is the number of data packets (slots) by frame;
- b_r is the number of bits assigned to the reference packets;
- b_p is the number of bits assigned to the preamble bits by frame;
- b_g is the number of equivalent bits that are corresponding to the guard interval;

The total number of bits per frame is:

$$b_{tot} = T_f R \tag{2.3}$$

where; - T_f is the total duration of the frame;

- R is the transmission rate of the date on the channel.

With these the efficiency of a TDMA frame can be determined by

$$\eta_F = \left(1 - \frac{b_{\sup}}{b_{tot}}\right) \times 100\% \tag{2.4}$$

In general is desired that the efficiency should not be less than 98 %, considering that it is computed including the bits introduced by source and data encoding, so the effective its value is significantly smaller.

A second parameter of a TDMA system is the ***number of channels*** that are offered by a TDMA system

$$N = \frac{m\left(B_{tot} - 2B_G\right)}{B_c} \tag{2.5}$$

where:

- m is the maximum number of users on each TDMA channel;
- B_{tot} is the total bandwidth assigned to the system;
- B_G are the 2 guard bandwidths that are on the ends of the occupied frequency domain;
- B_c is the bandwidth of a TDMA channel.

2.5. CHARACTERISTICS AND PARAMETERS OF CDMA SYSTEMS

In CDMA systems the information narrow-band signal is enlarged by combining it with a spreading code that has the code (*chip*) period smaller that the period of data by several measure orders. All the users are *using the same carrier frequency and the same frequency band simultaneously*, independently one from another, being individualized by the assigned spreading code. At receiver the transmitted data is compressed back into its bandwidth by *correlation* with its own code, while the rest of the signals, destined to other users are remaining of wideband , being treated as a noise from the demodulator point of view. Hence, in order to be able to detect the transmitted message, the receiver has to know the code used at the transmitter and to be perfect synchronized with it.

In a time-frequency-code coordination system, the CDMA allocation technique is shown in figure 2.7. The users are separated one another in code domain.

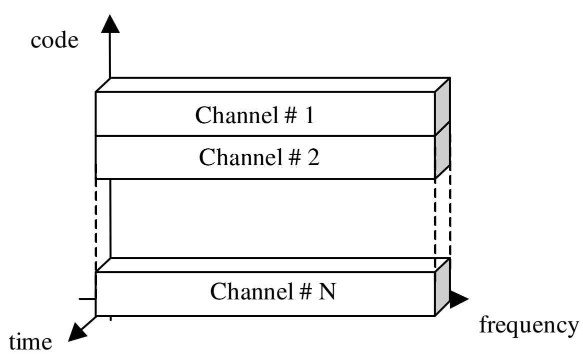

Fig. 2.7. Channel allocation in CDMA systems

The CDMA systems characteristics are:

- all the users of a CDMA system are using *the same carrier frequency and same frequency band in the same time.* To realize the duplexing one can use either FDD or TDD technique,;
- by contrast with TDMA and FDMA, the CDMA systems have a *soft capacity limit*. By increasing of the number of users, the value of the noise

level at reception will increase too, which leads to a degradation in performances for all users while their number is increasing;

- as the occupied bandwidth is *larger than the coherency bandwidth* of the channel, there will be an implicit frequency diversity in the transmitted signal that will combat *the selective frequency fading* effect;
- the transmitted *rate is very high* in the CDMA systems, so *the period of the spread data is very small*, much smaller than the temporal spreading of the channel. How the spreading sequence have very low correlation properties, the replicas that are late with more that one chip interval are appearing as noise at reception. For increasing the performances a *RAKE receiver* can be used that is combining the delayed replicas of the received signal in order to encounter the effect of multipath propagation;
- one of the problems that can appear in the CDMA systems is that of the *own jamming* that is caused by the fact that the used spreading frequencies are not perfectly orthogonal one another; because of that, at the de-correlation of a certain signal, there will appear and some certain contributions are due to other users;
- another problem that can appear is that of *caption of the receiver by another signa*l if its power is much higher than the desired signal.

The maximum number of users in a CDMA system is given by the number of orthogonal codes on a set. If pseudonoise (PN) codes are used, they will be generated using shift registers of length *n*; then the number of users will be

$$N = 2^n - 1 \tag{2.6}$$

The longer the shift register is, and, consequently the longer the code is, the number of users increases. On other hand a large value of n will increase the occupied bandwidth and will make more difficult the receiver synchronization. However, the maximum number of users is rarely reached, because of the soft limit of the capacity.

2.6. MULTIPLE ACCESS SYSTEMS USING PACKET RADIO (PR)

The Packet Radio (PR) technique assume that a number of users try to access simultaneously, in a non-coordinated manner, an unique radio channel in the high (1-30 MHz) or ultra-high (3-10GHz). frequency bandwidth. The transmission is made in packets of the same duration and modeled as a random process and, by contrast with the techniques described above, the users are not individualized in frequency, time or code. When two users transmits data simultaneously, a collision is produced and the data transmitted by both users are altered. To solve this problem, a set of rules had been established for transmitting the data forming a protocol; the widest known protocols are ALOHA and CSMA.

Those techniques have several advantages like they are easy to implement and do not require supplementary information, but they have also several drawbacks since they have a reduced spectral efficiency and if successive collisions appears the may induce considerable delays.

2.4.1. ALOHA

In the basic ALOHA system the following scenario is assumed: there are K users and each user transmits a packets of duration T_p, every NT_p seconds, where $N>>K$. If any of the K users wishes to transmit, it transmits immediately, using the whole available bandwidth. Note that the user *does not wait its "turn"* as in TDMA not is be assigned a narrow bandwidth as it FDMA. It must be noted that the ALOHA system assumes that the channel is, generally, not being used, so the average rate the packets from one user are entering the network is $1/NT_p$. In TDMA or FDMA system the packet rate that enters the network from one user is $1/KT_p$, significantly larger then in ALOHA.

One main problem faced when using ALOHA is that if two or more users transmit in the same time interval, at the receiver will arrive several packets simultaneously; since the packets not orthogonal in time (TDMA) or in frequency (FDMA), the receiver is and is unable to separate them, so none of them is received correctly. It is said a collision occurs and the packets have to be retransmitted after a given period of time.

If, the users rarely use the network the probability of a collision is small. To quantify this discussion we will assume that each user transmits at *an average rate* equal to $1/NT_p$, so the average rate considering all the packets that are entering the network is

$$\lambda = \frac{K}{NT_p} \text{ packets/s} \tag{2.7}$$

We shall also assume that all *users transmits independently one-another*, so that knowledge of the past history of what user u_j transmitted, and when, will yield no information concerning when the next packet will enter the network. That is, we assume that packets enter the network *randomly*. In this case the probability that m packets will enter the network in a time interval τ is given by the Poisson distribution

$$P(m,\tau) = \frac{e^{-\lambda\tau}(\lambda\tau)^m}{m!} \tag{2.8}$$

Thus, considering that user (1) transmits a packet at some time instant T_i. Then, referring to figure 1.8 we see that if any other user transmits a packet in the time interval

$$T_i - T_p < t < T_i + T_p \tag{2.9}$$

which is over an interval of duration $2T_p$ around T_i, there will be a "collision" and neither of the packets will be received correctly.

The probability that no other user is entering the channel over an interval $2T_p$ is found by setting $m = 0$ and $\tau = 2T_p$ in (2.8). The result is,

$$P_0 = P(m=0, \tau = 2T_p) = \frac{e^{\frac{-2K}{N}}\left(\frac{2K}{N}\right)^0}{0!} = e^{\frac{-2K}{N}} \tag{2.10}$$

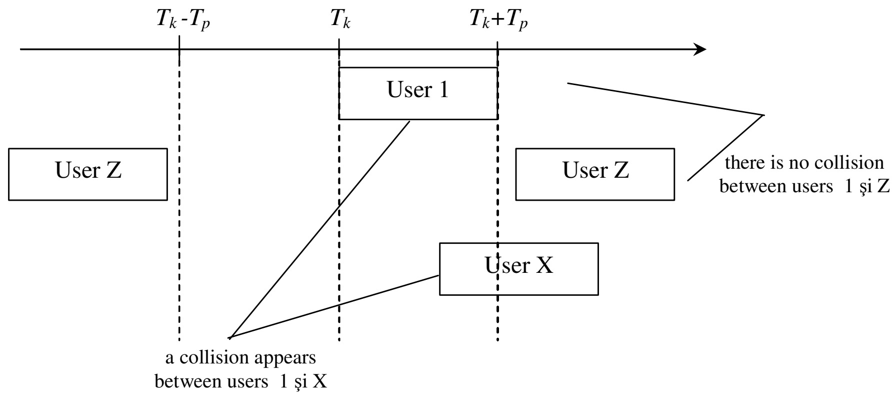

Fig. 1.8. The collision process in ALOHA

The average number of packets that are successfully received per second, is called the *throughput, S,* and is equal to

$$S = \frac{K}{NT_p} P_0 = \frac{K}{NT_p} e^{\frac{-2K}{N}} = \frac{\rho}{T_p} e^{-2\rho} \tag{2.11}$$

where $\rho = \dfrac{K}{N}$ is the density the packets are entering the network. The dependency of the throughput on the density the packets are entering the network is shown in figure 2.9.

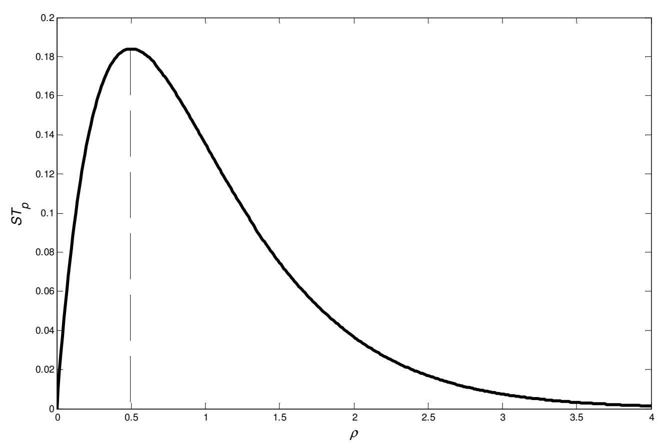

Fig. 2.9. The throughput vs. the density the packets are entering the network

It can be seen that the system achieves maximum throughput for

$$\rho = \frac{1}{2} \;\Rightarrow\; K = \frac{N}{2} \tag{2.12}$$

in which case the maximum throughput is

$$S_{\max} = \frac{1}{2T_p} e^{-1} \tag{2.13}$$

If the number of users increases over $N/2$, the number of collisions increases too and the throughput decreases dramatically. By contrast, in TDMA or FDMA the throughput is maximum when all the users transmit with the maximum rate $R_{\max} = W = 1/T_p$, that is

$$S_{\max,TDMA} = S_{\max,FDMA} = \frac{1}{T_p} \tag{2.14}$$

Comparing ALOHA with TDMA/FDMA with respect to throughput, we can see that

$$\frac{\left(S_{\max} T_p\right)_{ALOHA}}{\left(S_{\max} T_p\right)_{TDMA}} = \frac{1}{2e} = 0{,}1839 \tag{2.15}$$

which shows that the maximum throughput is 18,4 % with respect to the one achieved by FDMA/TDMA. This is one main disadvantages of the ALOHA system.

A second disadvantage is that there might be long delays between the transmission of a packet its correct reception, due to successive collisions between packets and the necessity of retransmitting them. If the distance between the transmitter and the receiver is d, the propagation time, is

$$T_d = \frac{d}{c} \tag{2.16}$$

where d is the distance traveled and c=3 10^8m/s is the speed of light.

Figure 2.10 illustrates the case where a collision occurs and retransmission is required. Note that packet is transmitted, and is received after a delay T_d. After examining the entire message an error was detected and a NAK message of duration T_n returned to each user. The system restrains each individual user from re-transmitting for random and independent wait time intervals T_w, so that hopefully, the users who collided before will not do so again. The time between the second reception and the first reception of packet i, is

$$\Delta = T_p + T_n + T_w + 2T_d \qquad (2.17)$$

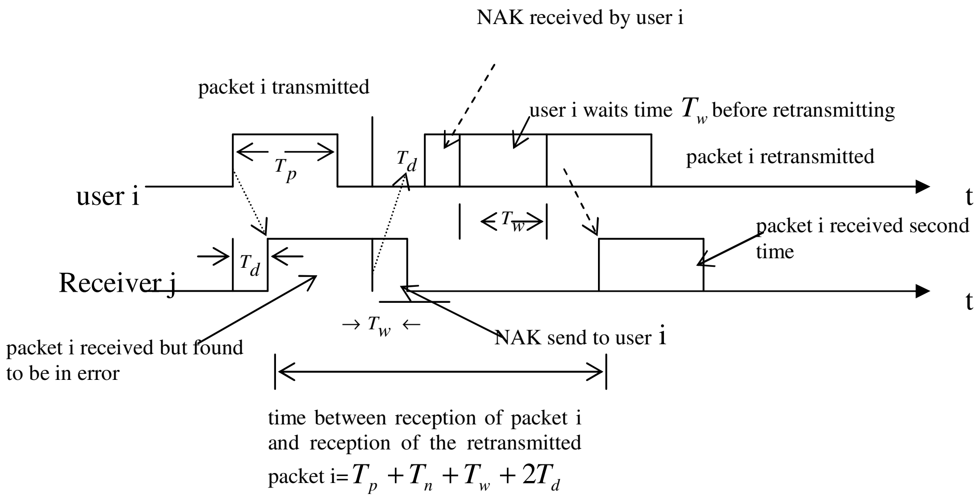

Fig.2.10. Collision occurrence in ALOHA

The average response delay D is given by

$$D = T_d P_0 + (T_d + \Delta)(1-P_0)P_0 + (T_d + 2\Delta)(1-P_0)^2 P_0 + \ldots\ldots + (T_d + q\Delta)(1-P_0)^q P_0 + \ldots\ldots \qquad (2.17)$$

where $T_d + q\Delta$ is the time elapsed between the transmission of a packet for the first time and its reception after being transmitted a total number of q+1 times; and $(1-P_0)^q P_0$ is the probability that a packet was received in error q times and receiver correctly the q+1th time. Using the well known identities

$$\begin{aligned} 1 + x + x^2 + \ \ldots \ &= \frac{1}{1-x} \\ 1 + 2x + 3x^2 + \ \ldots \ &= \frac{1}{(1-x)^2} \end{aligned} \qquad (2.18)$$

and making the appropriate mathematical manipulations, (2.17) becomes

$$D = T_d + \frac{\Delta}{P_0}(1-P_0) = T_d + \Delta\left(e^{2\rho} - 1\right) \qquad (2.18)$$

It can be observed that, when the throughput is maximum (that is ρ=0.5), the average delay is

$$D\,|_{C_{\max}} = T_d + \Delta(e-1) \tag{2.19}$$

2.4.2. SLOTTED ALOHA

The performances of ALOHA system can be increased by using the following simple protocol: we will divide the time scale into intervals each of duration T_p , named *time slots*, as shown in figure 2.11a. Then every user that wishes to transmit at time instance t, $T_i < t \le T_{i+1}$, must wait and transmit at time T_{i+1}; that is each user must *synchronize* its transmission to start at the beginning of a time slot. The protocol is named Slotted ALOHA.

As a result, collisions occur as shown in figure 2.12b only between packets wishing to transmit in a T_p second interval and not in $\boldsymbol{2}T_p$ second interval as in the pure ALOHA system described earlier.

The probability of a collision is obtained from Poisson distribution (2.10) with m=1 and τ=T_p,

$$P_0 = P(m=0; \tau = T_p) = e^{-\frac{K}{N}} = e^{-\rho} \tag{2.20}$$

where $\rho = K/N$ is the density the packets are entering the network and $\lambda = \rho / T_p$ is the average rate considering all the packets that are entering the network.

The throughput is determined, like for pure ALOHA by multiplying the density the packets are entering the network by the probability of their correct reception

$$S = \lambda P_0 = \frac{K}{NT_p} e^{-\frac{K}{N}} = \frac{\rho}{T_p} e^{-\rho} \tag{2.21}$$

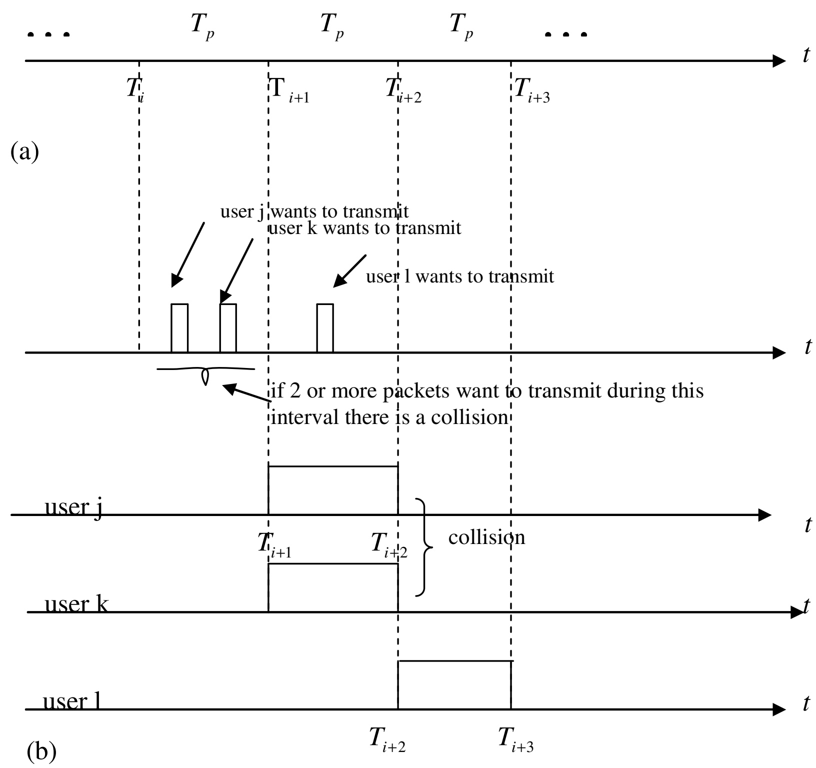

Fig.2.11. Collision occurrence in slotted ALOHA

The dependency of the throughput on the density of the packets is ilustrated in figure 2.12, in comparison with pure ALOHA.

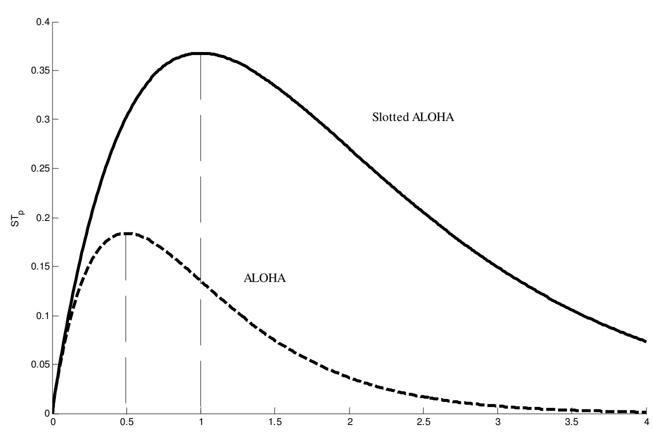

Fig.2.12 Throughput for Slotted ALOHA compared to ALOHA

We can easily determine the maximum value of the capacity, obtained for ρ=1

$$C_{\max,\,\mathrm{sinc}} = \frac{1}{T_p} e^{-1} \tag{2.22}$$

Comparing ALOHA with slotted ALOHA with respect to efficieny we have

$$\frac{C_{\max,\ \mathrm{sinc}} T_p}{C_{\max,\ \mathrm{nesinc}} T_p} = \frac{0{,}3688}{0{,}1839} = 2 \tag{2.23}$$

so the maximum efficiency obtained for slotted ALOHA is double the one obtained for pure ALOHA. Moreover, the value ρ=1 shows that, on average, for maximum throughput, the number of users is K=N. However, if K exceeds N, the number of collisions increase abruptly and the throughput decreases exponentially;

The average delay from the transmission of a packet and its correct reception can be determined similarly with the ALOHA case

$$D = T_d + \Delta(e^{\rho} - 1) \tag{2.24}$$

We must note that this delay is increasing slowly while the system load increases. When the throughput is maximum the delay is

$$D|_{C=\max} = T_d + \Delta(e - 1) \tag{2.23}$$

the result being the same as the one obtained in ALOHA case

One of the main disadvantages regarding the ALOHA protocol (pure or slotted) is that the system does not use any information regarding the data trafic on the transmission channel and of its level of occupance, which lead to a low value of the throughput, even in the slotted ALOHA case.

2.4.3. CARRIER SENSE MULTIPLE ACCESS (CSMA)

In the CSMA protocol the following scenario is assumed: any user that wishes to transmit data on the common channel has to determine first wether it is busy (active) or free (inactive). When a packet is received *busy* signal is transmitted. When a collision occurs, the receiver that detects the erronated packet transmitts a „*jam*" signal, on a common frequency f_c, received by all the users, announcing that the channel is bussy. In this case all the users will wait a random interval of time and then sense the channel again. The protocol is efficient when all the users are geografically clustered together, so the propagation delays is small. This situation is specific to Ethernet / LAN networks.

The system performances are evaluated, as in ALOHA type systems, by its throughput S and average time delay between the the time instant a packet is ready to be transmitted and the time when it is finally received.

In order to determine the system throughput, we will consider first the ideal case in which the propagation delay is zero and no collisions occurs. In this case, if all the users transmitts data of duraton T_p, the packets will succeed one another in order on the time axis as suggested in fig. 2.13; T_B is the waiting interval, moddeled as a random variable.

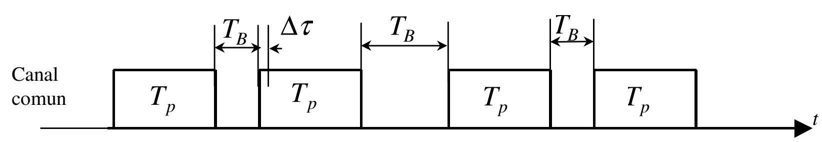

Fig. 2.13. Common channel time usage

The probability density of T_B is determined using the fact that the users are independent and the data entrance in the nework is modelled like a Poisson random variable. The probability that the T_B interval is inactive is given by

$$P(T_B) = P(m = 0; \tau = T_B)P(m = 1; \tau = \Delta t) \qquad (2.24)$$

where the first term sugests that in the time interval T_B no packet enters the network, and in the next time interval $\Delta\tau$ immediately after $\Delta\tau << T_B$ a packet is entering the network. Using the Poisson distribution (2.8) in (2.24) we have

$$P(T_B) = e^{-\lambda T_B}(\lambda\,\Delta\tau)e^{-\lambda\Delta\tau} = \lambda\,\Delta\tau\, e^{-\lambda(T_B+\Delta\tau)} \qquad (2.25)$$

With this, the probability density function is

$$f(T_B) = \frac{P(T_B)}{\Delta\tau} = \lambda e^{-\lambda T_B} = \frac{\rho}{T_p} e^{-\frac{\rho T_B}{T_p}} \tag{2.26}$$

where $\lambda = \rho / T_p$ is the average number of packets entering the network, and $\rho = K / N$ is the density the packets are entering the network. The mean value of the inactive interval is

$$\overline{T}_B = \int_0^\infty T_B\, f(T_B) dT_B = \frac{T_p}{\rho} \tag{2.27}$$

With this the system throughput (efficiency) is

$$S = \frac{1}{T_p + \overline{T}_B} = \frac{1}{T_p} \frac{1}{1 + \frac{1}{\rho}} = \frac{1}{T_p} \frac{\rho}{1 + \rho} \tag{2.28}$$

When the propagation delay is non-zero, the others with a certain delay, which might lead to collisions, receive the information transmitted by one user. We consider the following scenario: user (1) "listens" the channel at time instant t_1, finds it inactive and transmitts a packet of data immediately. The data is received at time $t_3 = t_1 + T_d$, when a busy signal is transmitted. If any other user listen the channel at time instant t_2, immediately following t_1 ($t_2 - t_1 < T_d$), it will find it inactive and transmit data, so a collision occurs. This situation is illustrated in figure 2.14.

The probability that no data packet is transmitted in the time interval $t \in (t_1,\, t_1 + T_d)$ can be obtained from the Poisson distribution (2.8) by setting m=0 and τ=T_d,

$$P_{nc} = P(m = 0; \tau = T_d) = e^{-\rho \frac{T_d}{T_p}} \tag{2.29}$$

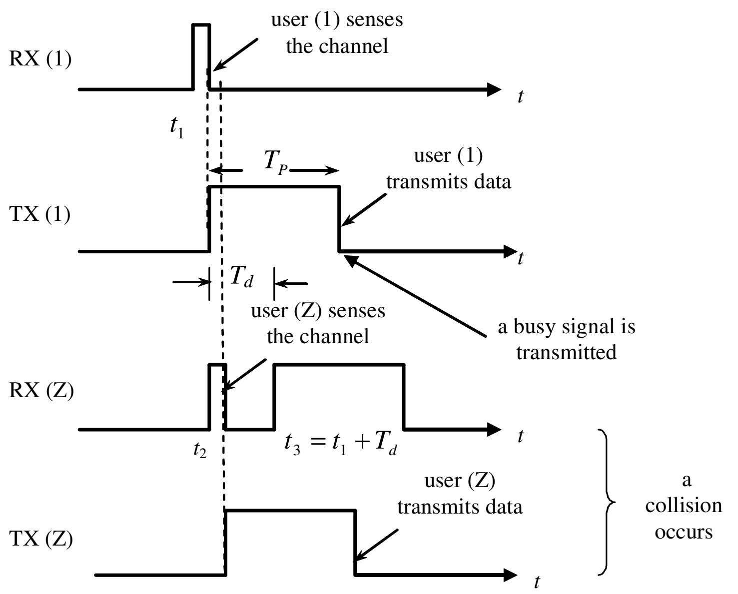

Fig. 2.14. Collision phenomenon in CSMA/CD

The probability of a collision is probabiliy that in this interval at least one packet enters the network

$$P_{nc} = P(m = 0; \tau = T_d) = e^{-\rho\frac{T_d}{T_p}} \tag{2.30}$$

To evaluate the collision effect on the system throughput we consider the scenario illustrated in figure 2.15.; by x ($0 \le x \le T_d$) we will denote the time interval between the beginning of a packet and that of a collision, which can be modelled as a random variable

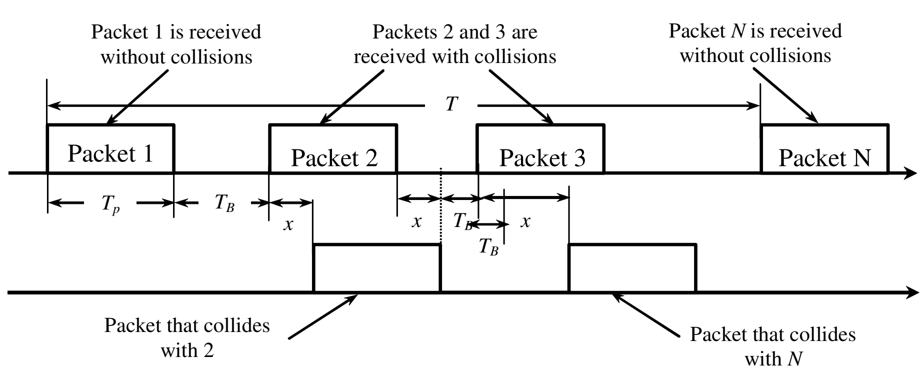

Fig. 2.15. Collision model in CSMA/CD

The average time between the receipt of two correct packets is

$$\overline{T} = \left(T_p + \overline{T}_B\right)P_{nc} + \left[2\left(T_p + \overline{T}_B\right) + \overline{x}\right]P_c P_{nc} + \left[3\left(T_p + \overline{T}_B\right) + 2\overline{x}\right]P_c^2 P_{nc} + \ \ldots \tag{2.31}$$

where the first term represents the time between the receipt of two correct packets if there is no collision times the probability of no collision, the second term is the time between the receipt of two correct packets if there is one collision at its first transmission and no collision at the second one, and so on. After algebraic manipulation we obtain

$$\overline{T} = \left(T_p + \overline{T}_B\right)\frac{1}{P_{nc}} + \overline{x}\,\frac{P_c}{P_{nc}} \tag{2.32}$$

Using (2.27), (2.29) and (2.30) and the fact that $0 \le x \le T_d$, we can find the superior and inferior limits of $\overline{T}$ and of the system throughput *S* as follows

- if x=0, $\Rightarrow$ $\overline{T}$ is minimum, *S* is maximum, so

$$S \le S_{\max} = \frac{\rho}{T_p}\,\frac{e^{-\rho\frac{T_d}{T_p}}}{1+\rho} \tag{2.33}$$

- if x=T_d, $\overline{T}$ is maximum, *S* is minimum, so

$$S \ge S_{\min} = \frac{\rho}{T_p}\,\frac{e^{-\rho\frac{T_d}{T_p}}}{1+\rho+\rho\dfrac{T_d}{T_p}\left(1-e^{-\rho\frac{T_d}{T_p}}\right)} \tag{2.34}$$

Note that when T_d / T_p decreases, the throughput increases assymptotically towards arround 0.9; when T_d / T_p decreases the throughput decreases also. We can conclude that the system is efficient only for small distances

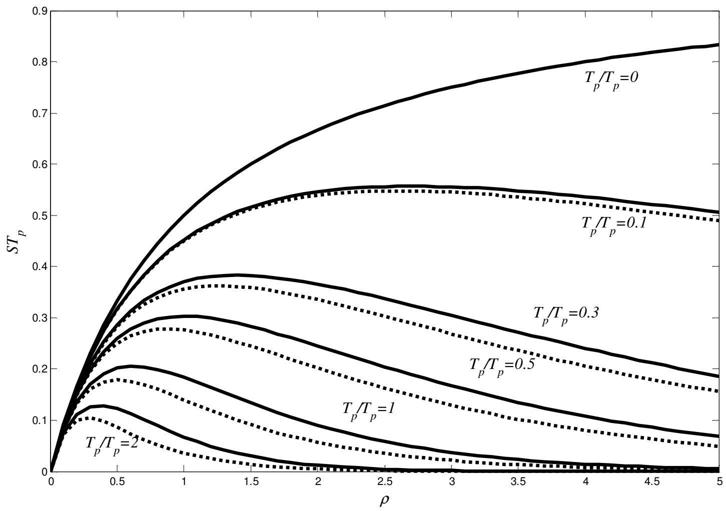

Fig. 2.16. CSMA efficiency as function of data density and $T_d\,/T_p$.

If we want to obtain for CSMA the same throughput for ALOHA

$$C_{\max,CSMA}=\frac{\rho}{T_p}\frac{e^{-\rho\frac{T_d}{T_p}}}{1+\rho}=C_{ALOHA}=\frac{\rho}{T_p}e^{-2\rho} \tag{1.50}$$

we have

$$\frac{T_d}{T_p}=2-\frac{\lg(1+\rho)}{\rho}\leq 2 \tag{1.51}$$

so CSMA is more efficient then ALOHA only if $T_d\,/T_p$>2.

Similarly, comparing CSMA with synchronized ALOHA we have

$$\frac{T_d}{T_p}=1-\frac{\lg(1+\rho)}{\rho}=1 \tag{1.52}$$

A value $T_d=T_p$ corresponds to a relatively large distance between users. For instance, if we assume that the users are transmitting data packets of 1000 bits with a rate of 10 Mb/s, we have $T_p=T_d=10^3/10^7=10^{-4}\,s$, which corresponds to a distance of $d=cT_d=310^8\,10^{-4}=30\,km$