Medium Access Control

This lesson introduces several Medium Access Control (MAC) algorithms which are specifically adapted to the wireless domain. Medium access control comprises all mechanisms that regulate user access to a medium using SDM, TDM, FDM, or CDM. MAC belongs to layer 2, the data link control layer (DLC). Layer 2 is subdivided into the logical link control (LLC), layer 2b, and the MAC, layer 2a. The task of DLC is to establish a reliable point to point or point to multi-point connection between different devices over a wired or wireless medium.

The lesson will explain why special MACs are needed in the wireless domain and why standard MAC schemes known from wired networks often fail. (In contrast to wired networks, hidden and exposed terminals or near and far terminals present serious problems here.) While SDM and FDM are typically used in a rather fixed manner, i.e., a certain space or frequency (or frequency hopping pattern) is assigned for a longer period of time; the main focus of this topic is on TDM mechanisms. TDM can be used in a very flexible way, as tuning in to a certain frequency does not present a problem, but time can be allocated on demand and in a distributed fashion. Well-known algorithms are Aloha (in several versions), different reservation schemes, or simple polling. Finally, the use of CDM is discussed again to show how a MAC scheme using CDM has to assign certain codes to allow the separation of different users in code space. It is of importance to note that one typically does not use a single scheme in its pure form but mixes schemes to benefit from the specific advantages.

Motivation for a specialized MAC

The main question in connection with MAC in the wireless is whether it is possible to use elaborated MAC schemes from wired networks, for example, CSMA/CD as used in the original specification of IEEE 802.3 networks (Ethernet).

Carrier Sense Multiple Access with Collision Detection, (CSMA/ CD) which works as follows. A sender senses the medium (a wire or Fiber cable) to see if it is free. If the medium is busy, the sender waits until it is free. If the medium is free, the sender starts transmitting data and continues to listen into the medium. If the sender detects a collision while sending, it stops at once and sends a jamming signal. Why does this scheme fail in wireless networks? CSMA/CD is not really interested in collisions at the sender, but rather in those at the receiver. The signal should reach the receiver without collisions. But the sender is the one detecting collisions. This is not a problem using a wire, as more or less the same signal strength can be assumed all over the wire if the length of the wire stays within certain often standardized limits. If a collision occurs somewhere in the wire, everybody will notice it. It does not matter if a sender listens into the medium to detect a collision at its own location while in reality is waiting to detect a possible collision at the receiver.

The situation is different in wireless networks. Remember, the strength of a signal decreases proportionally to the square of the distance to the sender. Obstacles attenuate the signal even further. The sender may now apply carrier sense and detect an idle medium. The sender starts sending – but a collision happens at the receiver due to a second sender (hidden terminal problem). The same can happen to the collision detection. The sender detects no collision and assumes that the data has been transmitted without errors, but a collision might actually have

destroyed the data at the receiver. Collision detection is very difficult in wireless scenarios as the transmission power in the area of the transmitting antenna is several magnitudes higher than the receiving power. So, this very common MAC scheme from wired network fails in a wireless scenario. The following sections show some more scenarios where schemes known from fixed networks fail.

Hidden and exposed terminals

Consider the scenario with three mobile phones as shown in Figure 1. The transmission range of A reaches B, but not C (the detection range does not reach C either). The transmission range of C reaches B, but not A. Finally, the transmission range of B reaches A and C, i.e., A cannot detect C and vice versa. A starts sending to B, C does not receive this transmission. C also wants to send something to B and senses the medium. The medium appears to be free, the carrier sense fails. C also starts sending causing a collision at B. But A cannot detect this collision at B and continues with its transmission. A is hidden from C and vice versa.

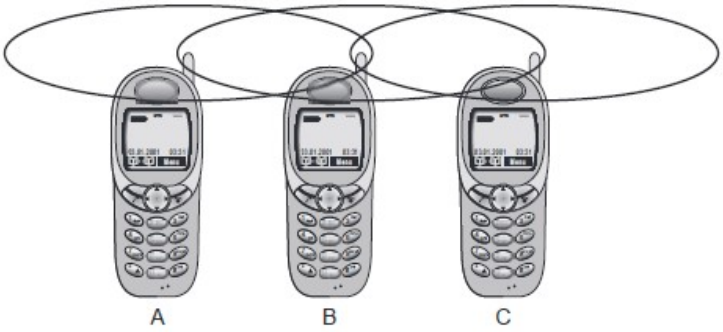

Figure 1 Hidden and Exposed terminals

While hidden terminals may cause collisions, the next effect only causes unnecessary delay. Now consider the situation where B sends something to A and C wants to transmit data to some other mobile phone outside the interference ranges of A and B. C senses the carrier and detects that the carrier is busy (B's signal). C postpones its transmission until it detects the medium as being idle again. But as A is outside the interference range of C, waiting is not necessary. Causing a 'collision' at B does not matter because the collision is too weak to propagate to A. In this situation, C is exposed to B.

Near and far terminals

Consider the situation as shown in Figure 2. Where A and B are both sending with the same transmission power. As the signal strength decreases proportionally to the square of the distance, B's signal drowns out A's signal. As a result, C cannot receive A's transmission. Now think of C as being an arbiter for sending rights (e.g., C acts as a base station coordinating media access). In this case, terminal B would already drown out terminal A on the physical layer. C in return would have no chance of applying a fair scheme as it would only hear B. The near/ far effect is a severe problem of wireless networks using CDM. All signals should arrive at the receiver with more or less the same strength. Otherwise a person standing closer to somebody could always speak louder than a person further away. Even if the senders were separated by code, the closest one would simply drown out the others. Precise power control is needed to receive all senders with the same strength at a receiver.

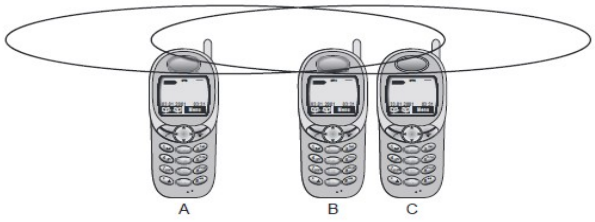

Figure 2 Near and Far Terminals

SDMA

Space Division Multiple Access (SDMA) is used for allocating a separated space to users in wireless networks. A typical application involves assigning an optimal base station to a mobile phone user. The mobile phone may receive several base stations with different quality. A MAC algorithm could now decide which base station is best, taking into account which frequencies (FDM), time slots (TDM) or code (CDM) are still available (depending on the technology). Typically, SDMA is never used in isolation but always in combination with one or more other schemes. The basis for the SDMA algorithm is formed by cells and sectorized antennas which constitute the infrastructure implementing space division multiplexing (SDM)

FDMA

Frequency Division Multiple Access (FDMA) comprises all algorithms allocating frequencies to transmission channels according to the frequency division multiplexing (FDM) scheme. Allocation can either be fixed (as for radio stations or the general planning and regulation of frequencies) or dynamic (i.e., demand driven). Channels can be assigned to the same frequency at all times, i.e., pure FDMA, or change frequencies according to a certain pattern, i.e., FDMA combined with TDMA. The latter example is the common practice for many wireless systems to circumvent narrowband interference at certain frequencies, known as frequency hopping. Sender and receiver have to agree on a hopping pattern, otherwise the receiver could not tune to the right frequency. Hopping patterns are typically fixed, at least for a longer period. The fact that it is not possible to arbitrarily jump in the frequency space (i.e., the receiver must be able to tune to the right frequency) is one of the main differences between FDM schemes and TDM schemes.

Furthermore, FDM is often used for simultaneous access to the medium by base station and mobile station in cellular networks. Here the two partners typically establish a duplex channel, i.e., a channel that allows for simultaneous transmission in both directions. The two directions, mobile station to base station and vice versa are now separated using different frequencies. This scheme is then called Frequency Division Duplex (FDD). Again, both partners have to know the

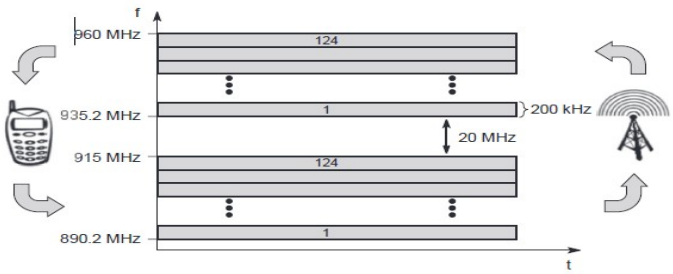

Figure 3 Frequency Division Multiplexing for Multiple Access and Duple

frequencies in advance; they cannot just listen into the medium. The two frequencies are also known as uplink, i.e., from mobile station to base station or from ground control to satellite, and as downlink, i.e., from base station to mobile station or from satellite to ground control.

For example, Figure 3 shows FDM and FDD in a mobile phone network based on the GSM standard for 900 MHz. The basic frequency allocation scheme for GSM is fixed and regulated by national authorities. (Certain variations may exist regarding the frequencies mentioned in the examples.) All uplinks use the band between 890.2 and 915 MHz, all downlinks use 935.2 to 960 MHz. According to FDMA, the base station, shown on the right side, allocates a certain frequency for up- and downlink to establish a duplex channel with a mobile phone. Up- and downlink have a fixed relation. If the uplink frequency is fu = 890 MHz + n·0.2 MHz, the downlink frequency is fd = fu + 45 MHz, i.e., fd = 935 MHz + n·0.2 MHz for a certain channel n. The base station selects the channel. Each channel (uplink and downlink) has a bandwidth of 200 kHz. This illustrates the use of FDM for multiple access (124 channels per direction are available at 900 MHz) and duplex according to a predetermined scheme.

TDMA

Compared to FDMA, Time Division Multiple Access (TDMA) offers a much more flexible scheme, which comprises all technologies that allocate certain time slots for communication, i.e., controlling TDM. Now tuning in to a certain frequency is not necessary, i.e., the receiver can stay at the same frequency the whole time. Using only one frequency, and thus very simple receivers and transmitters, many different algorithms exist to control medium access. As already mentioned, listening to different frequencies at the same time is quite difficult, but listening to many channels separated in time at the same frequency is simple. Almost all MAC schemes for wired networks work according to this principle, e.g., Ethernet, Token Ring, ATM etc.

Now synchronization between sender and receiver has to be achieved in the time domain. Again this can be done by using a fixed pattern similar to FDMA techniques, i.e., allocating a certain time slot for a channel, or by using a dynamic allocation scheme. Dynamic allocation schemes require an identification for each transmission as this is the case for typical wired MAC schemes (e.g., sender address) or the transmission has to be announced beforehand. MAC addresses are quite often used as identification. This enables a receiver in a broadcast medium to recognize if it really is the intended receiver of a message. Fixed schemes do not need identification, but are not as flexible considering varying bandwidth requirements. The following are several examples for fixed and dynamic schemes as used for wireless transmission. Typically, those schemes can be combined with FDMA to achieve even greater flexibility and transmission capacity.

Fixed TDM

The simplest algorithm for using TDM is allocating time slots for channels in a fixed pattern. This results in a fixed bandwidth and is the typical solution for wireless phone systems. MAC is quite simple, as the only crucial factor is accessing the reserved time slot at the right moment. If this synchronization is assured, each mobile station knows its turn and no interference will happen. The fixed pattern can be assigned by the base station, where competition between different mobile stations that want to access the medium is solved.

Fixed access patterns (at least fixed for some period in time) fit perfectly well for connections with a fixed bandwidth. Furthermore, these patterns guarantee a fixed delay – one can transmit, e.g., every 10 ms as this is the case for standard DECT systems. TDMA schemes with fixed access patterns are used for many digital mobile phone systems like IS-54, IS-136, GSM, DECT, PHS, and PACS.

Figure 4 shows how these fixed TDM patterns are used to implement multiple access and a duplex channel between a base station and mobile station. Assigning different slots for uplink and downlink using the same frequency is called Time Division Duplex (TDD). As shown in the figure, the base station uses one out of 12 slots for the downlink, whereas the mobile station uses one out of 12 different slots for the uplink. Uplink and downlink are separated in time. Up to 12 different mobile stations can use the same frequency without interference using this scheme. Each connection is allotted its own up- and downlink pair. In the example below, the pattern is repeated every 10 ms, i.e., each slot has a duration of 417 µs. This repetition guarantees access to the medium every 10 ms, independent of any other connections.

While the fixed access patterns, are perfectly apt for connections with a constant data rate (e.g., classical voice transmission with 32 or 64 kbit/s duplex), they are very inefficient for bursty data or asymmetric connections. If temporary bursts in data are sent from the base station to the mobile station often or vice versa (as in the case of web browsing, where no data transmission occurs while reading a page, whereas clicking on a hyperlink triggers a data transfer from the mobile station, often to the base station, often followed by huge amounts of data returned from the web server). This general scheme still wastes a lot of bandwidth. It is too static, too inflexible for data communication. In this case, connectionless, demand-oriented TDMA schemes can be used, as discussed below.

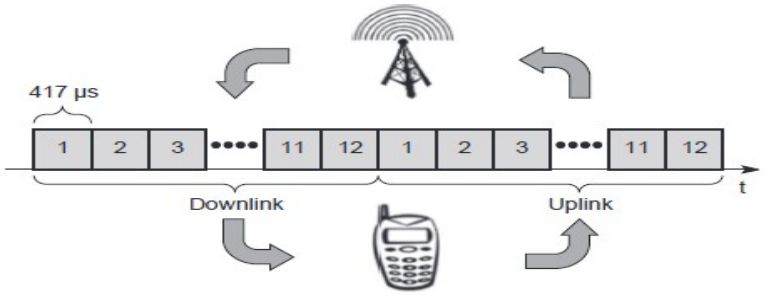

Figure 4 Time Division Multiplexing Access and Duplex

Classical Aloha

As mentioned above, TDMA comprises all mechanisms controlling medium access according to TDM. But what happens if TDM is applied without controlling access? This is exactly what the classical Aloha scheme does. Aloha neither coordinates medium access nor does it resolve contention on the MAC layer. Instead, each station can access the medium at any time as shown in Figure 5. This is a random access scheme, without a central arbiter controlling access and without coordination among the stations. If two or more stations access the medium at the same time, a collision occurs and the transmitted data is destroyed. Resolving this problem is left to higher layers (e.g., retransmission of data). The simple Aloha works fine for a light load and does not require any complicated access mechanisms.

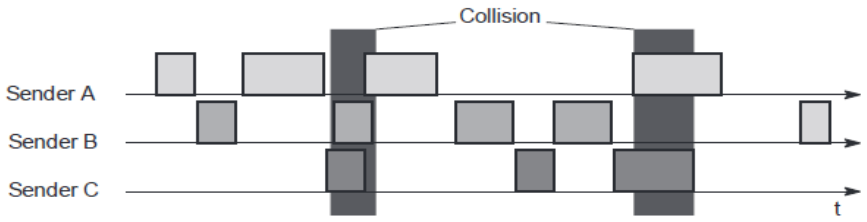

Figure 5 Classical Aloha Multiple Access

Slotted Aloha

The first refinement of the classical Aloha scheme is provided by the introduction of time slots (slotted Aloha). In this case, all senders have to be synchronized, transmission can only start at the beginning of a time slot as shown in Figure 6. Still, access is not coordinated. Under the assumption stated above, the introduction of slots raises the throughput from 18 per cent to 36 percent, i.e., slotting doubles the throughput. Both basic Aloha principles occur in many systems that implement distributed access to a medium. Aloha systems work perfectly well under a light load (as most schemes do), but they cannot give any hard transmission guarantees, such as maximum delay before accessing the medium, or minimum throughput. Here one needs additional mechanisms, e.g., combining fixed schemes and Aloha schemes. However, even new mobile communication systems like UMTS have to rely on slotted Aloha for medium access in certain situations (random access for initial connection set-up).

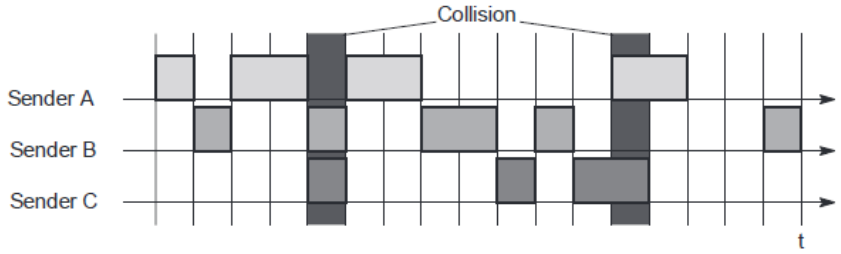

Figure 6 Slotted Aloha Multiple Access

Carrier Sense Multiple Access

One improvement to the basic Aloha is sensing the carrier before accessing the medium. This is what carrier sense multiple access (CSMA) schemes generally do. Sensing the carrier and accessing the medium only if the carrier is idle decreases the probability of a collision. But, as already mentioned in the introduction, hidden terminals cannot be detected, so, if a hidden terminal transmits at the same time as another sender, a collision might occur at the receiver. This basic scheme is still used in most wireless LANs

Several versions of CSMA exist. In non-persistent CSMA, stations sense the carrier and start sending immediately if the medium is idle. If the medium is busy, the station pauses a random amount of time before sensing the medium again and repeating this pattern. In p-persistent CSMA systems nodes also sense the medium, but only transmit with a probability of p, with the station deferring to the next slot with the probability 1-p, i.e., access is slotted in addition. In 1-persistent CSMA systems, all stations wishing to transmit access the medium at the same time,

as soon as it becomes idle. This will cause many collisions if many stations wish to send and block each other. To create some fairness for stations waiting for a longer time, back-off algorithms can be introduced, which are sensitive to waiting time as this is done for standard Ethernet.

CSMA with collision avoidance (CSMA/ CA) is one of the access schemes used in wireless LANs following the standard IEEE 802.11. Here sensing the carrier is combined with a back-off scheme in case of a busy medium to achieve some fairness among competing stations. Another, very elaborate scheme is elimination yield – non -preemptive multiple access (EY-NMPA). Here several phases of sensing the medium and accessing the medium for contention resolution are interleaved before one 'winner' can finally access the medium for data transmission. Here, priority schemes can be included to assure preference of certain stations with more important data.

Multiple Access with Collision Avoidance

One of the initial problems is that of hidden terminals. How do the previous access schemes solve this? To all schemes with central base stations assigning TDM patterns, the problem of hidden terminals is unknown. If the terminal is hidden for the base station it cannot communicate anyway. But as mentioned above, more or less fixed access patterns are not as flexible as Aloha schemes. What happens when no base station exists at all? This is the case in so-called ad-hoc networks.

Multiple access with collision avoidance (MACA) presents a simple scheme that solves the hidden terminal problem, does not need a base station, and is still a random access Aloha scheme – but with dynamic reservation. Figure 7 shows the same scenario as Figure 1 with the hidden terminals. Remember, A and C both want to send to B. A has already started the transmission, but is hidden for C, C also starts with its transmission, thereby causing a collision at B.

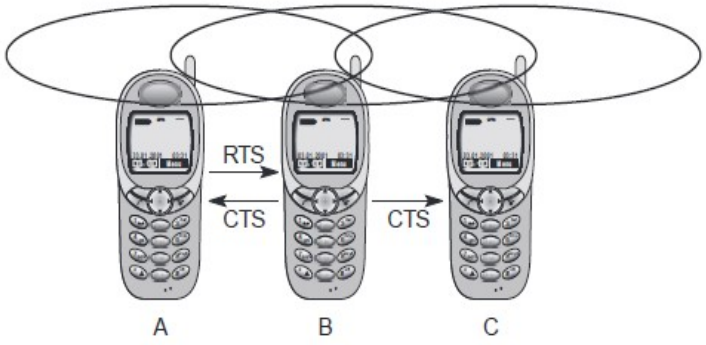

Figure 7 MACA can avoid hidden terminals

With MACA, A does not start its transmission at once, but sends a Request to Send (RTS) first. B receives the RTS that contains the name of sender and receiver, as well as the length of the future transmission. This RTS is not heard by C, but triggers an acknowledgement from B, called Clear to Send (CTS). The CTS again contains the names of sender (A) and receiver (B) of the user data, and the length of the future transmission. This CTS is now heard by C and the medium for future use by A is now reserved for the duration of the transmission. After receiving a CTS, C is not allowed to send anything for the duration indicated in the CTS toward B. A collision

cannot occur at B during data transmission, and the hidden terminal problem is solved – provided that the transmission conditions remain the same. What happens when another station moves into the transmission range of B after the transmission of CTS.

Still, collisions can occur during the sending of an RTS. Both A and C could send an RTS that collides at B. RTS is very small compared to the data transmission, so the probability of a collision is much lower. B resolves this contention and acknowledges only one station in the CTS (if it was able to recover the RTS at all). No transmission is allowed without an appropriate CTS. This is one of the medium access schemes that is optionally used in the standard IEEE 802.11.

Can MACA also help to solve the 'exposed terminal' problem? Remember, B wants to send data to A, C to someone else. But C is polite enough to sense the medium before transmitting, sensing a busy medium caused by the transmission from B. C defers, although C could never cause a collision at A. With MACA, B has to transmit an RTS first as shown in Figure 8 below containing the name of the receiver (A) and the sender (B). C does not react to this message as it is not the receiver, but A acknowledges using a CTS which identifies B as the sender and A as the receiver of the following data transmission. C does not receive this CTS and concludes that A is outside the detection range. C can start its transmission assuming it will not cause a collision at A. The problem with exposed terminals is solved without fixed access patterns or a base station. One problem of MACA is clearly the overheads associated with the RTS and CTS transmissions for short and time-critical data packets, this is not negligible. MACA assumes symmetrical

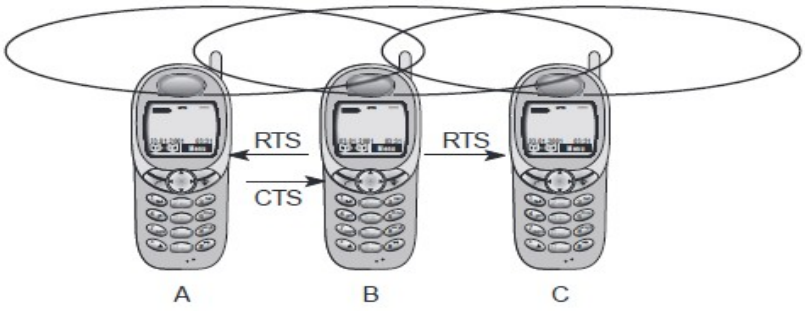

Figure 8 MACA can avoid exposed terminals

transmission and reception conditions. Otherwise, a strong sender, directed antennas etc. could counteract the above scheme.

Figure 9 shows simplified state machines for a sender and receiver. The sender is idle until a user requests the transmission of a data packet. The sender then issues an RTS and waits for the right to send. If the receiver gets an RTS and is in an idle state, it sends back a CTS and waits for data. The sender receives the CTS and sends the data. Otherwise, the sender would send an RTS again after a time-out (e.g., the RTS could be lost or collided). After transmission of the data, the sender waits for a positive acknowledgement to return into an idle state. The receiver sends back a positive acknowledgement if the received data was correct. If not, or if the waiting time for data is too long, the receiver returns into idle state. If the sender does not receive any acknowledgement or a negative acknowledgement, it sends an RTS and again waits for the right

to send. Alternatively, a receiver could indicate that it is currently busy via a separate RxBusy. Real implementations have to add more states and transitions, e.g., to limit the number of retries.

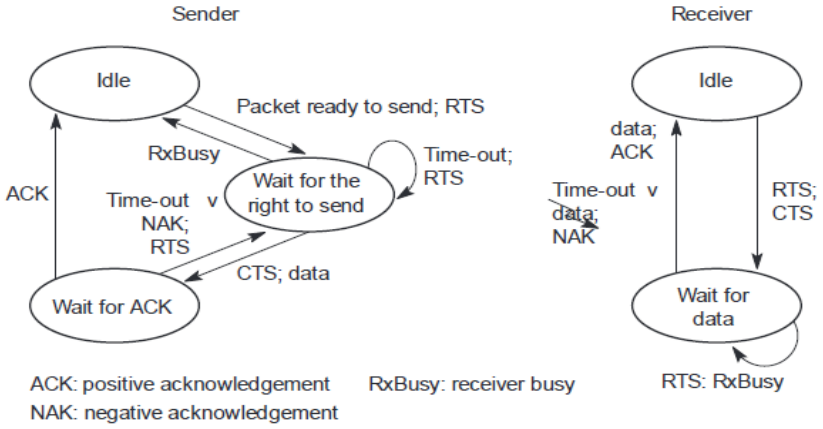

Figure 9 Protocol Machines for Multiple Access with Collision Avoidance

Polling

Where one station is to be heard by all others (e.g., the base station of a mobile phone network or any other dedicated station), Polling schemes can be applied. Polling is a strictly centralized scheme with one master station and several slave stations. The master can poll the slaves according to many schemes: round robin (only efficient if traffic patterns are similar over all stations), randomly, according to reservations. The master could also establish a list of stations wishing to transmit during a contention phase. After this phase, the station polls each station on the list. Similar schemes are used, e.g., in the Bluetooth wireless LAN and as one possible access function in IEEE 802.11 systems.

Inhibit Sense Multiple Access

Another combination of different schemes is represented by inhibit sense multiple access (ISMA). This scheme, which is used for the packet data transmission service Cellular Digital Packet Data (CDPD) in the AMPS mobile phone system, is also known as Digital Sense Multiple Access (DSMA). Here, the base station only signals a busy medium via a busy tone (called BUSY/IDLE indicator) on the downlink, refer to Figure 10. After the busy tone stops, accessing the uplink is not coordinated any further. The base station acknowledges successful transmissions, a mobile station detects a collision only via the missing positive acknowledgement. In case of collisions, additional back-off and retransmission mechanisms are implemented.

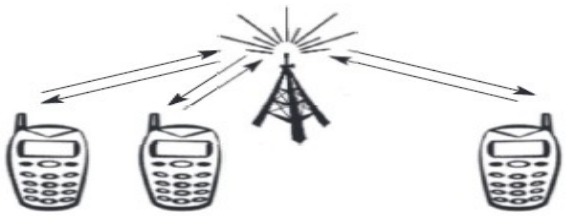

Figure 10 Inhibit Sense Multiple Access Using a Busy Tone