MOBILE COMPUTING NOTES

UNIT I

INTRODUCTION TO MOBILE COMMUNICATIONS

Introduction:

There are two different kinds of mobility: **user mobility** and **device portability**. User mobility refers to a user who has access to the same or similar telecommunication services at different places, i.e., the user can be mobile, and the services will follow him or her.

With device portability, the communication device moves (with or without a user). Many mechanisms in the network and inside the device have to make sure that communication is still possible while the device is moving. A typical example for systems supporting device portability is the mobile phone system, where the system itself hands the device from one radio transmitter (also called a base station) to the next if the signal becomes too weak.

With regard to devices, the term wireless is used. This only describes the way of accessing a network or other communication partners, i.e., without a wire. The wire is replaced by the transmission of electromagnetic waves through 'the air' (although wireless transmission does not need any medium).

A communication device can thus exhibit one of the following characteristics:

● **Fixed and wired**: This configuration describes the typical desktop computer in an office. Neither weight nor power consumption of the devices allow for mobile usage. The devices use fixed networks for performance reasons.

● **Mobile and wired**: Many of today's laptops fall into this category; users carry the laptop from one hotel to the next, reconnecting to the company's network via the telephone network and a modem.

● **Fixed and wireless**: This mode is used for installing networks, e.g., in historical buildings to avoid damage by installing wires, or at trade shows to ensure fast network setup. Another example is bridging the last mile to a customer by a new operator that has no wired infrastructure and does not want to lease lines from a competitor.

● **Mobile and wireless**: This is the most interesting case. No cable restricts the user, who can roam between different wireless networks. Most technologies discussed in this book deal with this type of device and the networks supporting them. Today's most successful example for this category is GSM with more than 800 million users.

Applications

1. Vehicles

Today's cars already comprise some, but tomorrow's cars will comprise many wireless communication systems and mobility aware applications. Music, news, road conditions, weather reports, and other broadcast information are received via digital audio broadcasting (DAB) with 1.5 Mbit/s. For personal communication, a universal mobile telecommunications system (UMTS) phone might be available offering voice and data connectivity with 384 Kbit/s. For remote areas, satellite communication can be used, while the current position of the car is determined via the global positioning system (GPS). Cars driving in the same area build a local ad-hoc network for the fast exchange of information in emergency situations or to help each other keep a safe distance. In case of an accident, not only will the airbag be triggered, but the police and ambulance service will be informed via an emergency call to a service provider. Cars with this technology are already available. In the future, cars will also inform other cars about accidents via the ad-hoc network to help them slowdown in time, even before a driver can recognize an accident. Buses, trucks, and trains are already transmitting maintenance and logistic information to their home base, which helps to improve organization (fleet management), and saves time and money.

2. **Emergencies**

Just imagine the possibilities of an ambulance with a high-quality wireless connection to a hospital. Vital information about injured persons can be sent to the hospital from the scene of the accident. All the necessary steps for this particular type of accident can be prepared and specialists can be consulted for an early diagnosis. Wireless networks are the only means of communication in the case of natural disasters such as hurricanes or earthquakes. In the worst cases, only decentralized, wireless ad-hoc networks survive. The breakdown of all cabling not only implies the failure of the standard wired telephone system, but also the crash of all mobile phone systems requiring base stations.

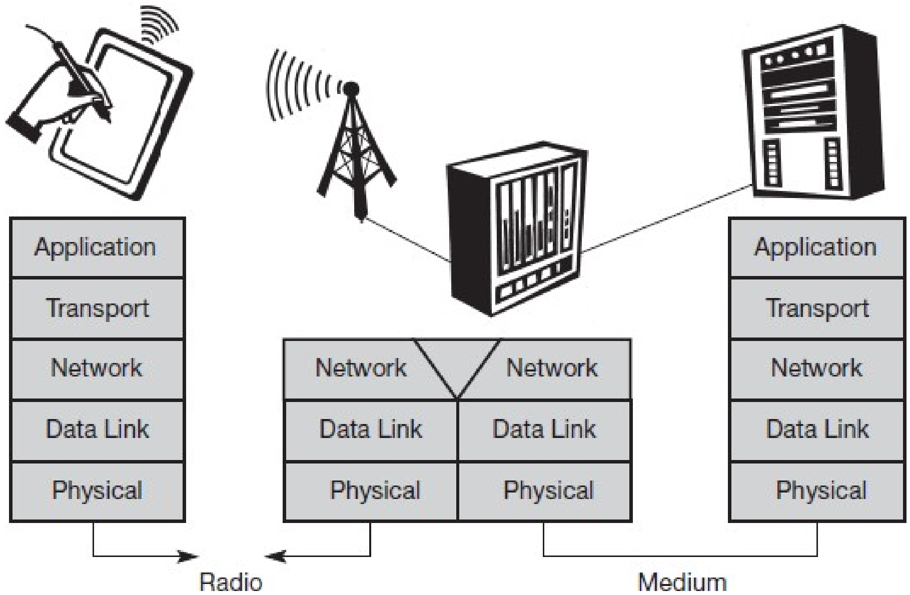
3. **Business**

A travelling salesman today needs instant access to the company's database: to ensure that files on his or her laptop reflect the current situation, to enable the company to keep track of all activities of their travelling employees, to keep databases consistent etc. With wireless access, the laptop can be turned into a true mobile office, but efficient and powerful synchronization mechanisms are needed to ensure data consistency. At home, the laptop connects via a WLAN or LAN and DSL to the Internet. Leaving home requires a handover to another technology, e.g., to an enhanced version of GSM, as soon as the WLAN coverage ends. Due to interference and other factors discussed in chapter 2, data rates drop while cruising at higher speed. Gas stations may offer WLAN hot spots as well as gas. Trains already offer support for wireless connectivity. Several more handovers to different technologies might be necessary before reaching the office. No matter when and where, mobile communications should always offer as good connectivity as possible to the internet, the company's intranet, or the telephone network.

4. **Replacement of wired networks**

In some cases, wireless networks can also be used to replace wired networks, e.g., remote sensors, for tradeshows, or in historic buildings. Due to economic reasons, it is often impossible to wire remote sensors for weather forecasts, earthquake detection, or to provide environmental information. Wireless connections, e.g., via satellite, can help in this situation. Tradeshows need a highly dynamic infrastructure, but cabling takes a long time and frequently proves to be too inflexible. Many computer fairs use WLANs as a replacement for cabling. Other cases for wireless networks are computers, sensors, or information displays in historical buildings, where excess cabling may destroy valuable walls or floors. Wireless access points in a corner of the room can represent a solution.

A simplified reference model



● **Physical layer**: This is the lowest layer in a communication system and is responsible for the conversion of a stream of bits into signals that can be transmitted on the sender side. The physical layer of the receiver then transforms the signals back into a bit stream. For wireless communication, the physical layer is responsible for frequency selection, generation of the carrier frequency, signal detection (although heavy interference may disturb the signal), modulation of data onto a carrier frequency and (depending on the transmission scheme) encryption.

● **Data link layer**: The main tasks of this layer include accessing the medium, multiplexing of different data streams, correction of transmission errors, and synchronization (i.e., detection of a data frame). Altogether, the data link layer is responsible for a reliable point-to point connection between two devices or a point-to-multipoint connection between one sender and several receivers.

● **Network layer**: This third layer is responsible for routing packets through a network or establishing a connection between two entities over many other intermediate systems. Important topics are addressing, routing, device location, and handover between different networks.

● **Transport layer**: This layer is used in the reference model to establish an end-to-end connection. Topics like quality of service, flow and congestion control are relevant, especially if the transport protocols known from the Internet, TCP and UDP, are to be used over a wireless link.

● **Application layer**: Finally, the applications (complemented by additional layers that can support applications) are situated on top of all transmission oriented layers. Topics of interest in this context are service location, support for multimedia applications, adaptive applications that can handle the large variations in transmission characteristics, and wireless access to the World Wide Web using a portable device. Very demanding applications are video (high data rate) and interactive gaming (low jitter, low latency).

Frequencies for radio transmission

Radio transmission can take place using many different frequency bands. Each frequency band exhibits certain advantages and disadvantages The figure shows frequencies starting at 300 Hz and going up to over 300 THz. Directly coupled to the frequency is the wavelength λ via the equation: $\lambda = c/f$, where $c \cong 3\cdot 108$ m/s (the speed of light in vacuum) and f the frequency. For traditional wired networks, frequencies of up to several hundred kHz are used for distances up to some km with twisted pair copper wires, while frequencies of several hundred MHz are used with coaxial cable (new coding schemes work with several hundred MHz even with twisted pair copper wires over distances of some 100 m). Fiber optics are used for frequency ranges of several hundred THz, but here one typically refers to the wavelength which is, e.g., 1500 nm, 1350 nm etc. (infra red). Radio transmission starts at several kHz, the very low frequency (VLF) range. These are very long waves. Waves in the low frequency (LF) range are used by submarines, because they can penetrate water and can follow the earth's surface. Some radio stations still use these frequencies, e.g., between 148.5 kHz and 283.5 kHz in Germany. The medium frequency (MF) and high frequency (HF) ranges are typical for transmission of hundreds of radio stations either as amplitude modulation (AM) between 520 kHz and 1605.5 kHz, as short wave (SW) between 5.9 MHz and 26.1 MHz, or as frequency modulation (FM) between 87.5 MHz and 108 MHz. The frequencies limiting these ranges are typically fixed by national regulation and, vary from country to country. Short waves are typically used for (amateur) radio transmission around the world, enabled by reflection at the ionosphere. Transmit power is up to 500 kW – which is quite high compared to the 1 W of a mobile phone. digital audio broadcasting (DAB) takes place as well (223–230 MHz and 1452–1472 MHz) and digital TV is planned or currently being installed (470– 862 MHz), reusing some of the old frequencies for analog TV. UHF is also used for mobile phones with analog technology (450–465 MHz), the digital GSM (890–960 MHz, 1710–1880 MHz), digital cordless telephones following the DECT standard (1880–1900 MHz), 3G cellular systems following the UMTS standard (1900–1980 MHz, 2020–2025 MHz, 2110–2190 MHz) and many more. VHF and especially UHF allow for small antennas and relatively reliable connections for mobile telephony. Super high frequencies (SHF) are typically used for directed microwave links (approx. 2–40 GHz) and fixed satellite services in the C-band (4 and 6 GHz), Ku-band (11 and 14 GHz), or Ka-band (19 and 29 GHz). Some systems are planned in the extremely high frequency (EHF) range which comes close to infra red. All radio frequencies are regulated to avoid interference, e.g., the German regulation covers 9 kHz–275 GHz. The next step into higher frequencies involves optical transmission, which is not only used for fiber optical links but also for wireless communications. Infra red (IR) transmission is used for directed links, e.g., to connect different buildings via laser links. The most widespread IR technology, infra red data association (IrDA), uses wavelengths of approximately 850–900 nm to connect laptops, PDAs etc. Finally, visible light has been used for wireless transmission for thousands of years. While light is not very reliable due to interference, but it is nevertheless useful due to built-in human receivers.

Data and Signals

To be transmitted, data must be transformed to electromagnetic signals
Both data and the signals that represent them can be either analog or digital in form.

Analog and Digital Data

Data can be analog or digital. The term analog data refers to information that is continuous; digital data refers to information that has discrete states
Analog data, such as the sounds made by a human voice, take on continuous values. When someone speaks, an analog wave is created in the air. This can be captured by a microphone and converted to an analog signal or sampled and converted to a digital signal. Digital data take on discrete values. For example, data are stored in computer memory in the form of O's and 1's. They can be converted to a digital signal or modulated into an analog signal for transmission across a medium.
Signals can be analog or digital. Analog signals can have an infinite number of values in a range; digital signals can have only a limited number of values

Antennas

As the name wireless already indicates, this communication mode involves 'getting rid' of wires and transmitting signals through space without guidance. We do not need any 'medium' (such as an ether) for the transport of electromagnetic waves.
Antennas couple electromagnetic energy to and from space to and from a wire or coaxial cable (or any other appropriate conductor). A theoretical reference antenna is the isotropic radiator, a point in space radiating equal power in all directions, i.e., all points with equal power are located on a sphere with the antenna as its center. The radiation pattern is symmetric in all directions two dimensional

Real antennas

Exhibit directive effects, i.e., the intensity of radiation is not the same in all directions from the antenna. The simplest real antenna is a thin, center-fed dipole, also called Hertzian dipole. The dipole consists of two collinear conductors of equal length, separated by a small feeding gap. The length of the dipole is not arbitrary, but, for example, half the wavelength λ of the signal to transmit results in a very efficient radiation of the energy. If mounted on the roof of a car, the length of $\lambda/4$ is efficient. This is also known as Marconi antenna. A $\lambda/2$ dipole has a uniform or omni-directional radiation pattern in one plane.

Signal propagation

Like wired networks, wireless communication networks also have senders and receivers of signals. However, in connection with signal propagation, these two networks exhibit considerable differences. In wireless networks, the signal has no wire to determine the direction of propagation, whereas signals in wired networks only travel along the wire (which can be twisted pair copper wires, a coax cable, but also a fiber etc.). As long as the wire is not interrupted or damaged, it typically exhibits the same characteristics at each point. One can precisely determine the behavior of a signal travelling along this wire, e.g., received power depending on the length. For wireless transmission, this predictable behavior is only valid in a vacuum, i.e., without matter between the sender and the receiver.

Transmission range: Within a certain radius of the sender transmission is possible, i.e., a receiver receives the signals with an error rate low enough to be able to communicate and can also act as sender.

● Detection range: Within a second radius, detection of the transmission is possible, i.e., the transmitted power is large enough to differ from background noise. However, the error rate is too high to establish communication.

● Interference range: Within a third even larger radius, the sender may interfere with other transmission by adding to the background noise. A receiver will not be able to detect the signals, but the signals may disturb other signals.

MULTIPLEXING

Bandwidth utilization is the wise use of available bandwidth to achieve specific goals. Efficiency can be achieved by multiplexing; privacy and anti-jamming can be achieved by spreading.

Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single data link. As data and telecommunications use increases, so does traffic. We can accommodate this increase by continuing to add individual links each time a new channel is needed; or we can install higher-bandwidth links and use each to carry multiple signals.

There are three basic multiplexing techniques: frequency-division multiplexing, wavelength-division multiplexing, and time-division multiplexing. The first two are techniques designed for analog signals, the third, for digital signals.

FDM is an analog multiplexing technique that combines analog signals.

WDM is an analog multiplexing technique to combine optical signals

TDM is a digital multiplexing technique for combining several low-rate channels into one high-rate one.

In synchronous TDM, the data rate of the link is *n* times faster, and the unit duration is *n* times shorter.

Space division multiplexing(SDM)

For wireless communication, multiplexing can be carried out in four dimensions: space, time, frequency, and code. In this field, the task of multiplexing is to assign space, time, frequency, and code to each communication channel with a minimum of interference and a maximum of medium utilization. The term communication channel here only refers to an association of sender(s) and receiver(s) who want to exchange data six channels ki and introduces a three dimensional coordinate system. This system shows the dimensions of code c, time t and frequency f. For this first type of multiplexing, space division multiplexing (SDM) The channels k1 to k3 can be mapped onto the three ‘spaces’ s1 to s3 which clearly separate the channels and prevent the interference ranges from overlapping. The space between the interference ranges is sometimes called guard space. Such a guard space is needed in all four multiplexing schemes presented.

Frequency division multiplexing(FDM)

Frequency division multiplexing (FDM) describes schemes to subdivide the frequency dimension into several non-overlapping frequency bands
. Each channel ki is now allotted its own frequency band as indicated. Senders using a certain frequency band can use this band continuously. Again, guard spaces are needed to avoid frequency band overlapping (also called adjacent channel interference). This scheme is used for radio stations within the same region, where each radio station has its own frequency. This very simple multiplexing scheme does not need complex coordination between sender and receiver: the receiver only has to tune in to the specific sender.

Time division multiplexing(TDM)

A more flexible multiplexing scheme for typical mobile communications is time division multiplexing (TDM). Here a channel ki is given the whole bandwidth for a certain amount of time, i.e., all senders use the same frequency but at different points in time .Again, guard spaces, which now represent time gaps, have to separate the different periods when the senders use the medium. In our highway example, this would refer to the gap between two cars. If two transmissions overlap in time, this is called co-channel interference. (In the highway example, interference between two cars results in an accident.) To avoid this type of interference, precise synchronization between different senders is necessary. This is clearly a disadvantage, as all senders need precise clocks or, alternatively, a way has to be found to distribute a synchronization signal to all senders. For a receiver tuning in to a sender this does not just involve adjusting the frequency, but involves listening at exactly the right point in time. However, this scheme is quite flexible as one can assign more sending time to senders with a heavy load and less to those with a light load. requency and time division multiplexing can be combined, i.e., a channel ki can use a certain frequency band for a certain amount of time Now guard spaces are needed both in the time and in the frequency dimension. This scheme is more robust against frequency selective interference, i.e., interference in a certain small frequency band. A channel may use this band only for a short period of time. Additionally, this scheme provides some (weak) protection against tapping, as in this case the sequence of frequencies a sender uses has to be known to listen in to a channel. The mobile phone standard GSM uses this combination of frequency and time division multiplexing for transmission between a mobile phone and a so-called base station

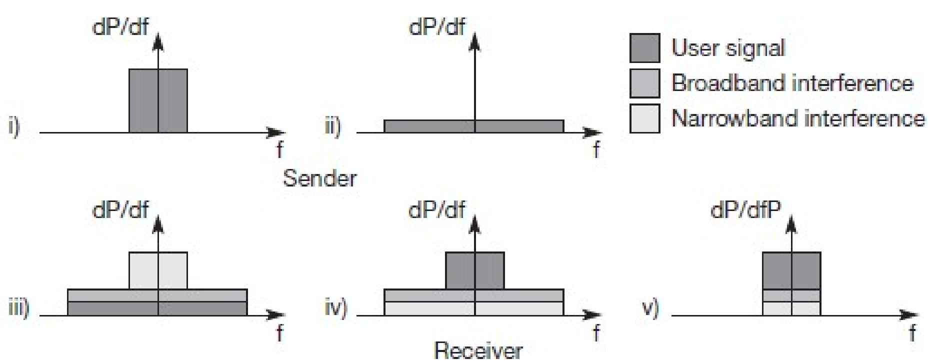
Code division multiplexing (CDM)

While SDM and FDM are well known from the early days of radio transmission and TDM is used in connection with many applications, code division multiplexing (CDM) is a relatively new scheme in commercial communication systems. First used in military applications due to its inherent security features (together with spread spectrum techniques, it now features in many civil wireless transmission scenarios thanks to the availability of cheap processing power channels ki use the same frequency at the same time for transmission. Separation is now achieved by assigning each channel its own 'code', guard spaces are realized by using codes with the necessary 'distance' in code space, e.g., orthogonal codes. The main advantage of CDM for wireless transmission is that it gives good protection against interference and tapping. Different codes have to be assigned, but code space is huge compared to the frequency space. Assigning individual codes to each sender does not usually cause problems. The main disadvantage of this scheme is the relatively high complexity of the receiver. A receiver has to know the code and must separate the channel with user data from the background noise composed of other signals and environmental noise. Additionally, a receiver must be precisely synchronized with the transmitter to apply the decoding correctly. The voice example also gives a hint to another problem of CDM receivers. All signals should reach a receiver with almost equal strength, otherwise some signals could drain others. If some people close to a receiver talk very loudly the language does not matter. The receiver cannot listen to any other person. To apply CDM, precise power control is required.

SPREAD SPECTRUM

In spread spectrum, we also combine signals from different sources to fit into a larger bandwidth, but our goals are somewhat different. Spread spectrum is designed to be used in wireless applications (LANs and WANs). In these types of applications, we have some concerns that outweigh bandwidth efficiency. In wireless applications, all stations use air (or a vacuum) as the medium for communicatiospread spectrum techniques add redundancy; they spread the original spectrum needed for each station.

Figure 2.32
Spread spectrum: spreading and despreading



There are two techniques to spread the bandwidth: frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS).

Frequency Hopping Spread Spectrum (FHSS)

The frequency hopping spread spectrum (FHSS) technique uses *M* different carrier frequencies that are modulated by the source signal. At one moment, the signal modulates one carrier frequency; at the next moment, the signal modulates another carrier frequency. Although the modulation is done using one carrier frequency at a time, *M* frequencies are used in the long run.

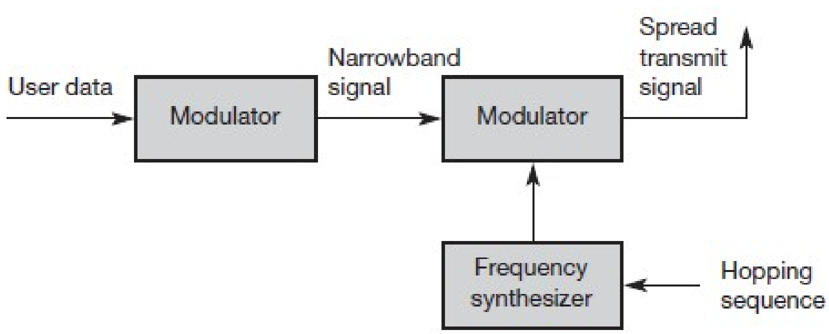


Figure 2.39
FHSS transmitter

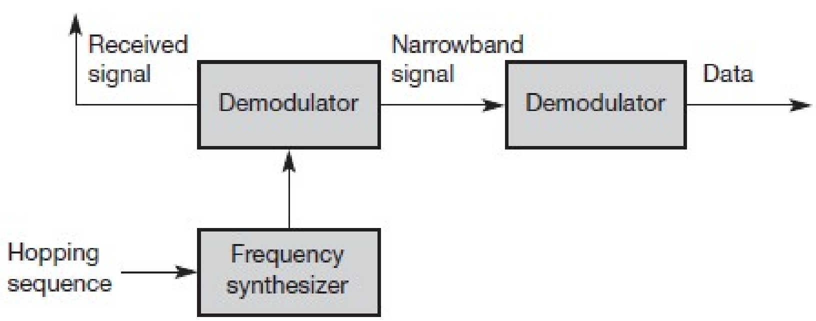


Figure 2.40
FHSS receiver

A pseudorandom code generator, called pseudorandom noise (PN), creates a k-bit pattern for every hopping period *Th•* The frequency table uses the pattern to find the frequency to be used for this hopping period and passes it to the frequency synthesizer. The frequency synthesizer creates a carrier signal of that frequency, and the source signal modulates the carrier signal.

Direct Sequence Spread Spectrum

The direct sequence spread spectrum (nSSS) technique also expands the bandwidth of the original signal, but the process is different. In DSSS, we replace each data bit with 11 bits using a spreading code. In other words, each bit is assigned a code of 11 bits, called chips, where the chip rate is 11 times that of the data bit. spreading code is 11 chips having the pattern 10110111000 (in this case). If the original signal rate is *N,* the rate of the spread signal is *lIN.* This means that the required bandwidth for the spread signal is 11 times larger than the bandwidth of the original signal. The spread signal can provide privacy if the intruder does not know the code. It can also provide immunity against interference if each station uses a different code.

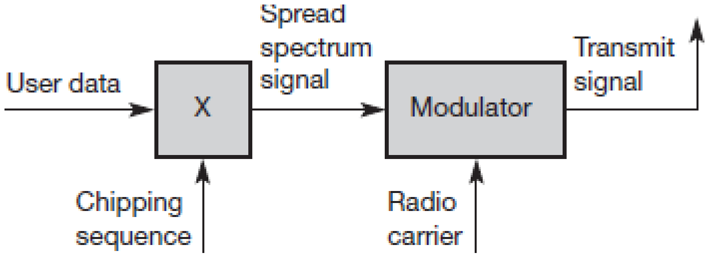
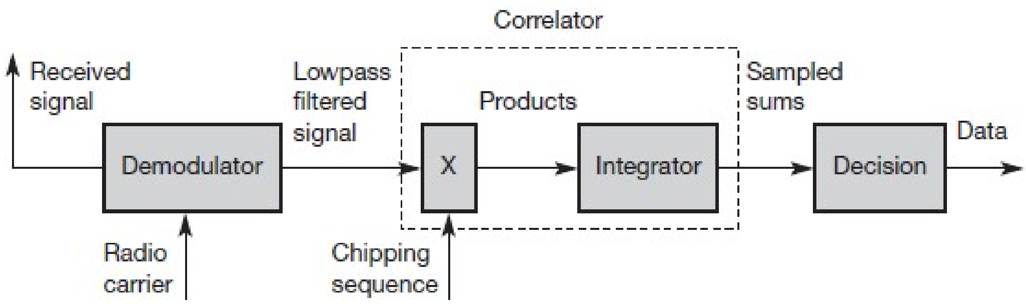


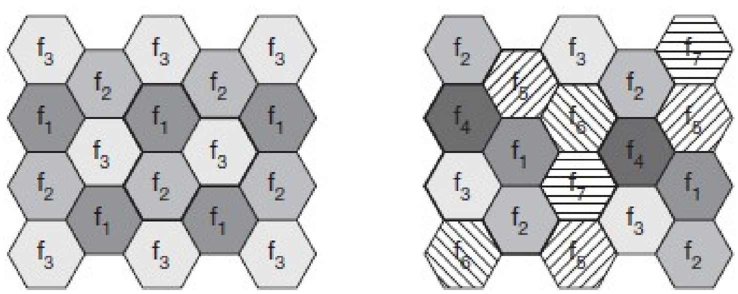
Figure 2.37
DSSS receiver



Cellular systems

Cellular systems for mobile communications implement SDM. Each transmitter,typically called a base station, covers a certain area, a cell. Cell radii can varyfrom tens of meters in buildings, and hundreds of meters in cities, up to tens ofkilometers in the countryside. The shape of cells are never perfect circles orhexagons but depend on the environment (buildings,mountains, valleys etc.), on weather conditions, and sometimes even on systemload. Typical systems using this approach are mobile telecommunication systems.

Figure 2.41
Cellular system with three and seven cell clusters



Advantages of cellular systems with small cells are the following:

● **Higher capacity**: Implementing SDM allows frequency reuse. If one transmitter is far away from another, i.e., outside the interference range, it can reuse the same frequencies. As most mobile phone systems assign frequencies to certain users (or certain hopping patterns), this frequency is blocked for other users. But frequencies are a scarce resource and, the number of concurrent users per cell is very limited. Huge cells do not allow for more users. On the contrary, they are limited to less possible users per km2. This is also the reason for using very small cells in cities where many more people use mobile phones.

● **Less transmission power**: While power aspects are not a big problem for base stations, they are indeed problematic for mobile stations. A receiver far away from a base station would need much more transmit power than the current few Watts. But energy is a serious problem for mobile handheld devices.

● **Local interference only**: Having long distances between sender and receiver results in even more interference problems. With small cells, mobile stations and base stations only have to deal with 'local' interference.

● **Robustness**: Cellular systems are decentralized and so, more robust against the failure of single components. If one antenna fails, this only influences communication within a small area.

MEDIUM ACCESS CONTROL(MAC)

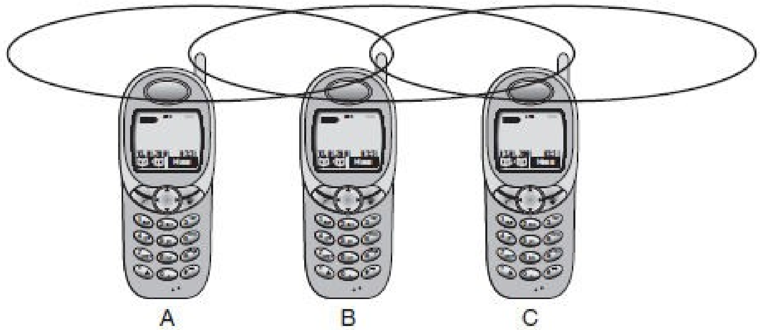
Motivation for a specialized MAC

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.

Consider carrier sense multiple access with collision detection, (CSMA/CD) which works as follows. A sender senses the medium (a wire or coaxial 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.

This scheme fails in wireless networks. 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. 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.

Hidden and exposed terminals

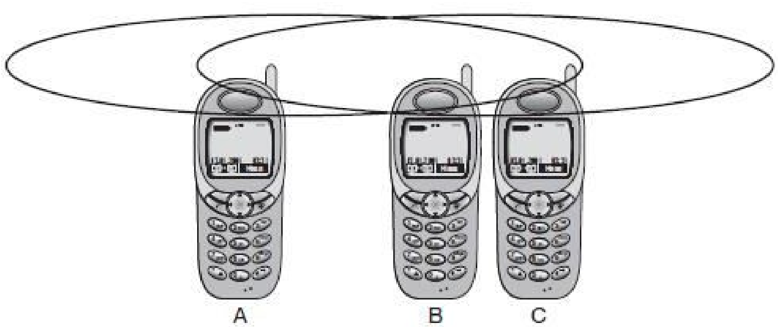


Consider the scenario with three mobile phones as shown in Figure. The transmission range of A reaches B, but not C (the detection range does not reach C either). The transmission range of C reaches B, but not A. Finally, the transmission range of B reaches A and C, i.e., A cannot detect C and vice versa.

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** for C and vice versa.

While hidden terminals may cause collisions, the next effect only causes unnecessary delay. Now consider the situation that 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. 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

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. The basis for the SDMA algorithm is formed by cells and sectorized antennas which constitute the infrastructure implementing space division multiplexing (SDM).

For wireless communication, multiplexing can be carried out in four dimensions: space, time, frequency, and code. In this field, the task of multiplexing is to assign space, time, frequency, and code to each communication channel with a minimum of interference and a maximum of medium utilization.

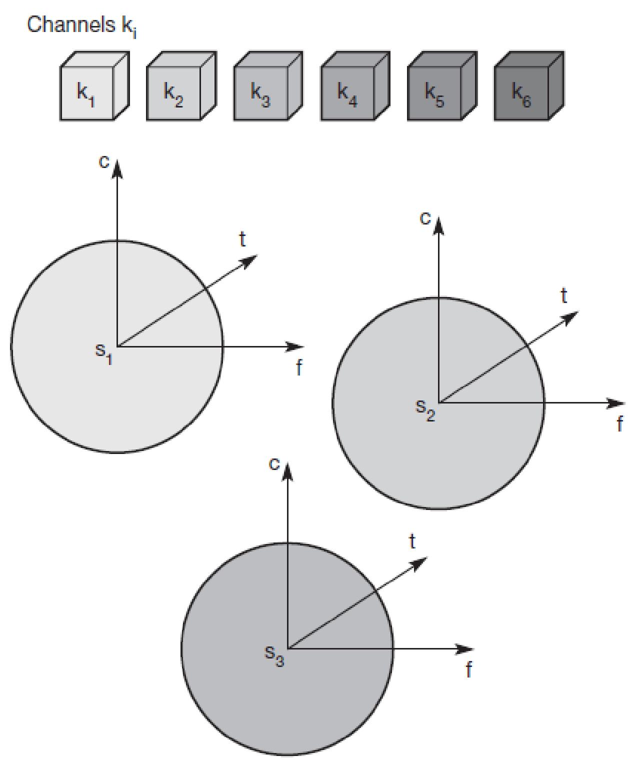


Figure shows six channels ki and introduces a three dimensional coordinate system. This system shows the dimensions of code c, time t and frequency f. For this first type of multiplexing,

space division multiplexing (SDM), the (three dimensional) space s_i is also shown. Here space is represented via circles indicating the interference range as introduced in Figure. The channels k1 to k3 can be mapped onto the three 'spaces' s1 to s3 which clearly separate the channels and prevent the interference ranges from overlapping. The space between the interference ranges is sometimes called **guard space**. Such a guard space is needed in all four multiplexing schemes presented.

For the remaining channels (k4 to k6) three additional spaces would be needed.

FDMA

Frequency division multiple access (FDMA) comprises all algorithms allocating frequencies to transmission channels according to the **frequency division multiplexing (FDM)**. 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. However, this scheme also has disadvantages. While radio stations broadcast 24 hours a day, mobile

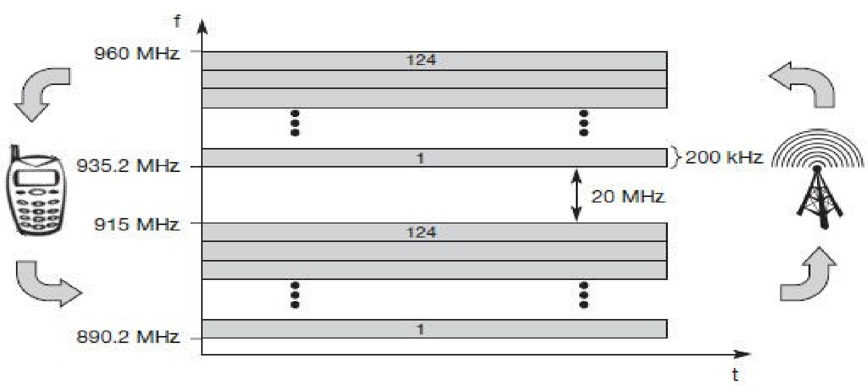


Figure 3.3
Frequency division multiplexing for multiple access and duplex

communication typically takes place for only a few minutes at a time. Assigning a separate frequency for each possible communication scenario would be a tremendous waste of (scarce) frequency resources. Additionally, the fixed assignment of a frequency to a sender makes the scheme very inflexible and limits the number of senders.

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. 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. 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.

It is too static, too inflexible for data communication. In this case, connectionless, demand-oriented TDMA schemes can be used.

Classical Aloha

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, a scheme which was invented at the University of Hawaii and was used in the ALOHANET for wireless connection of several stations. Aloha neither coordinates medium access nor does it resolve contention on the MAC layer. Instead, each station can access the medium at any time. 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.

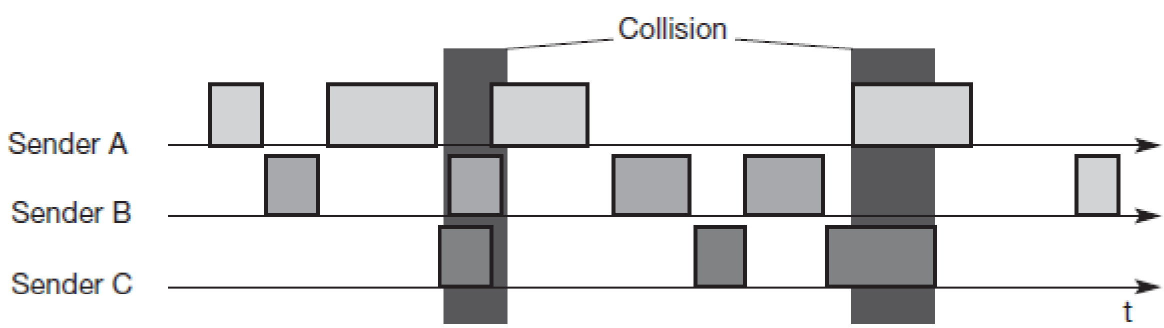


Fig: Classical Aloha multiple access

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. Still, access is not coordinated.

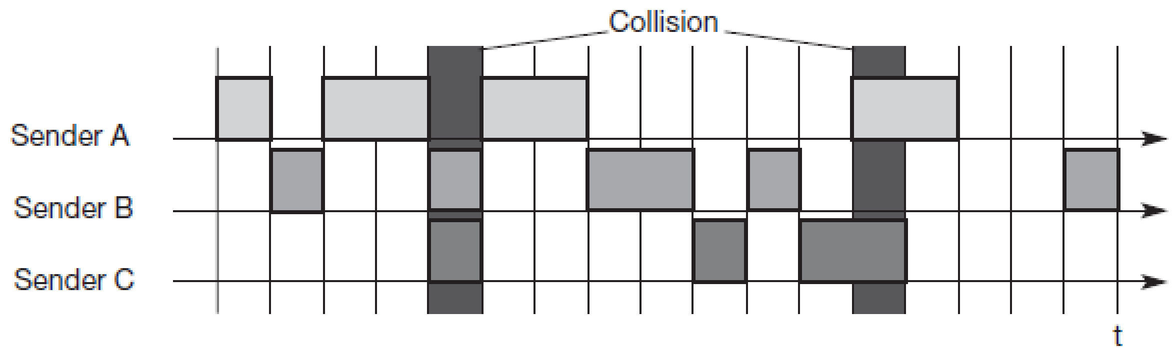


Fig: Slotted Aloha multiple access

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.

Carrier sense multiple access(CSMA)

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. 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.

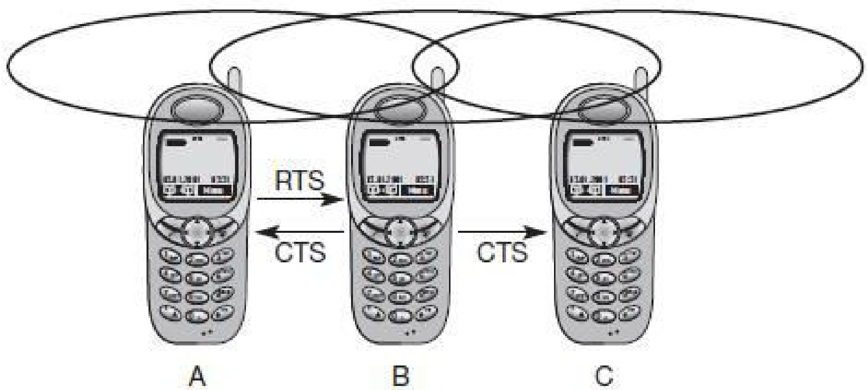
Several versions of CSMA exist. In **non-persistent CSMA**, stations sense the carrier and start sending immediately if the medium is idle. If the medium is busy, the station pauses a random amount of time before sensing the medium again and repeating this pattern. In **p-persistent CSMA** systems nodes 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.

Demand assigned multiple access

A general improvement of Aloha access systems can also be achieved by **reservation** mechanisms and combinations with some (fixed) TDM patterns. These schemes typically have a reservation period followed by a transmission period. During the reservation period, stations can reserve future slots in the transmission period. While, depending on the scheme, collisions may occur during the reservation period, the transmission period can then be accessed without collision. Alternatively, the transmission period can be split into periods with and without collision. In general, these schemes cause a higher delay under a light load (first the reservation has to take place), but allow higher throughput due to less collisions.

One basic scheme is **demand assigned multiple access (DAMA)** also called **reservation Aloha**, a scheme typical for satellite systems. DAMA has two modes. During a contention phase following the slotted Aloha scheme, all stations can try to reserve future slots. For example, different stations on earth try to reserve access time for satellite transmission. Collisions during the reservation phase do not destroy data transmission, but only the short requests for data transmission. If successful, a time slot in the future is reserved, and no other station is allowed to transmit during this slot. Therefore, the satellite collects all successful requests (the others are destroyed) and sends back a reservation list indicating access rights for future slots. All ground stations have to obey this list. To maintain the fixed TDM pattern of reservation and transmission, the stations have to be synchronized from time to time. DAMA is an **explicit reservation** scheme. Each transmission slot has to be reserved explicitly.

Multiple access with collision avoidance

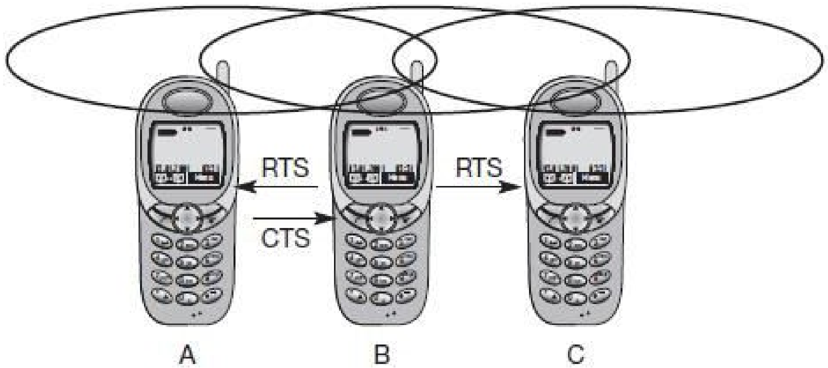


MACA can avoid hidden terminals

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 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.

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. (Another station could move 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 appropriate CTS.



MACA can avoid exposed terminals

MACA also help to solve the ‘exposed terminal’ problem. 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 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 these 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.

Comparison of S/T/F/CDMA

The table shows the MAC schemes without combination with other schemes. However, in real systems, the MAC schemes always occur in combinations. A very typical combination is constituted by SDMA/TDMA/FDMA as used in IS-54, GSM, DECT, PHS, and PACS phone systems, or the Iridium and ICO satellite systems. CDMA together with SDMA is used in the IS-95 mobile phone system and the Globalstar satellite system.

Approach	SDMA	TDMA	FDMA	CDMA
Idea	Segment space into cells/sectors	Segment sending time into disjoint time-slots, demand driven or fixed patterns	Segment the frequency band into disjoint sub-bands	Spread the spectrum using orthogonal codes
Terminals	Only one terminal can be active in one cell/one sector	All terminals are active for short periods of time on the same frequency	Every terminal has its own frequency, uninterrupted	All terminals can be active at the same place at the same moment, uninterrupted
Signal separation	Cell structure directed antennas	Synchronization in the time domain	Filtering in the frequency domain	Code plus special receivers
Advantages	Very simple, increases capacity per km^2	Established, fully digital, very flexible	Simple, established, robust	Flexible, less planning needed, soft handover
Disadvantages	Inflexible, antennas typically fixed	Guard space needed (multi-path propagation), synchronization difficult	Inflexible, frequencies are a scarce resource	Complex receivers, needs more complicated power control for senders
Comment	Only in combination with TDMA, FDMA or CDMA useful	Standard in fixed networks, together with FDMA/SDMA used in many mobile networks	Typically combined with TDMA (frequency hopping patterns) and SDMA (frequency reuse)	Used in many 3G systems, higher complexity, lowered expectations; integrated with TDMA/FDMA