The history of the wir\eless communications started with the understanding of magnetic and electric properties observed during the early days by the Chinese, Roman and Greek cultures and experiments carried out in the 17th and 18th centuries. A short history of wireless communication is presented in the tabular form:

Year	Description
1880	Hertz-Radio Communication
1897	Marconi- Radio Transmission
1933	FCC (Federal Communication Commission)
1938	FCC rules for regular services
1946	Bell telephone laboratories 52 MHz
1956	FCC - 450MHz (Simplex)
1964	Bell telephone active research 800 MHz
1964	FCC - 450 MHz (Full Duplex)
1969	FCC - 40 MHz bandwidth
1981	FCC ? release of cellular land phone in the 40 MHz
1982	At & T divested and Server RBOC (Regional Bell Operation Companies) formed to manage the cellular operation.
1984	Most RBOC market in operations
1986	FCC allocates 5MHz extended band.
1988	TDMA voted as digital cellular standard in North America.
1992	GSM (Group Special Mobile) operable Germany D2 system.
1993	CDMA (Code Division Multiple Access)
1994	PDCC (Personal Digital Cellular Operable) in Tokyo, Japan
1995	CDMA operable in Hong Kong
1996	Six Broad Band PCS (Personal Communication Services) licensed bands (120 MHz) almost reader 20 billion US dollar
1997	Broad band CDMA constructed and of the 3rd generation mobile.
1999	Powerful WLAN systems were evolved, such as Bluetooth. This uses 2.4 MHz spectrum.

Generations of Wireless Communication

1G

- This is the first generation of wireless telephone technology, mobile telecommunications, which was launched in Japan by NTT in 1979.
- The main technological development in this generation that distinguished the First Generation mobile phones from the previous generation was the use of multiple cell sites, and the ability to transfer calls from one site to the next site as the user travelled between cells during a conversation.
- It uses analog signals.
- It allows the voice calls in one country.

Cons:

- Poor quality of voice
- Poor life of Battery
- Size of phone was very large
- No security
- Capacity was limited
- Poor handoff reliability

2G

- This is the second generation of mobile telecommunication was launched in Finland in 1991.
- It was based on GSM standard.
- It enables data transmission like as text messaging (SMS - Short Message Service), transfer or photos or pictures (MMS ? Multimedia Messaging Service), but not videos.
- The later versions of this generation, which were called 2.5G using GPRS (General Packet Radio Service) and 2.75G using EDGE (Enhanced data rates for GSM Evolution) networks.
- It provides better quality and capacity.

Disadvantages

- Unable to handle complex data such as Video
- Requires strong digital signals

3G

- 3G is the third generation was introduced in early 2000s.
- The transmission of data was increased up to 2Mbits/s, which allows you to sending or receiving large email messages.
- The main difference between 3G and 2G is the use of packet switching rather than circuit switching for data transmission.
- Faster communication
- High speed web or more security
- Video conferencing
- 3D gaming
- TV streaming, Mobile TV, phone calls etc. are the features of 3G.

Disadvantages

- Costly
- Requirement of high bandwidth

- Expensive 3G phones
- Size of cell phones was very large.

4G

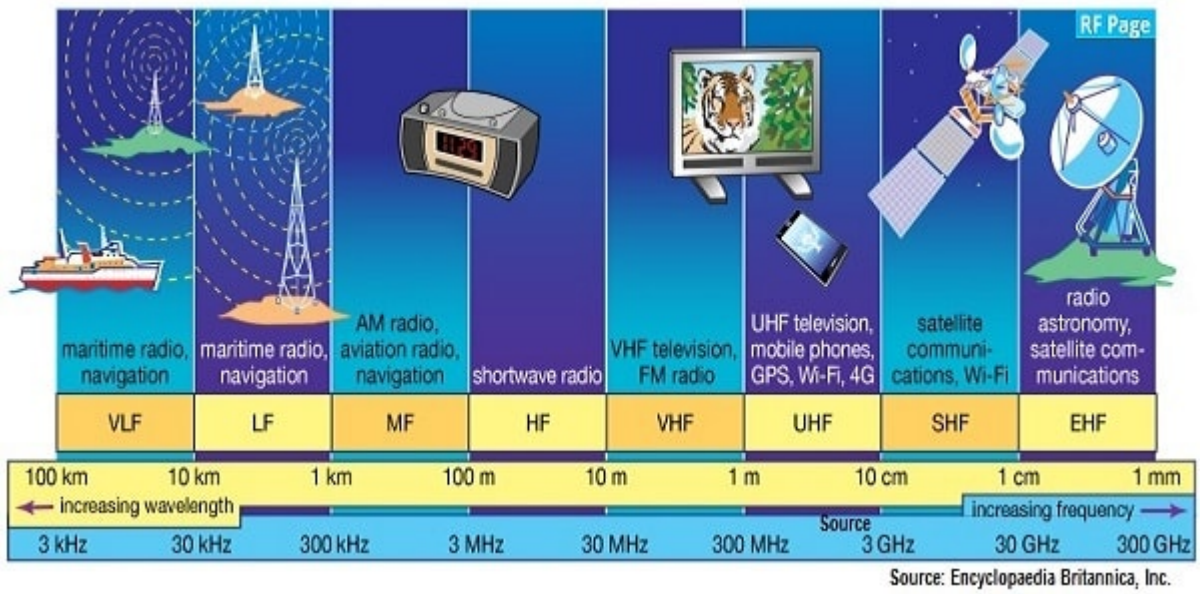
- 4G is the fourth generation of mobile telecommunication which was appeared in 2010.
- It was based on LTE (Long Term Evolution) and LTE advanced standards.
- Offer a range of communication services like video calling, real time language translation and video voice mail.
- It was capable of providing 100 Mbps to 1Gbps speed.
- High QoS (Quality of Service) and High security.
- The basic term used to describe 4G technology is MAGIC. Where : M - Mobile multiedia A - Anytime anywhere G - Global mobility support I - Integarted wireless solution C - Customized personal service

Disadvantages

- Uses more battery
- Difficult to implement
- Expensive equipment are required

5G

- It is refered to fifth generation wireless connection which will be probably implemented by 2020, or even some years earlier.
- Machine to machine communication can be possible in 5G.
- 5G will be able to performs Internet of Things (IoT) for smart home and smart city, connected cars etc.
- This generation will be based on lower cost, low battery consumption and lower latency than 4G equipment.
- There will be much fater transmission rate of data to the previous versions. Thus the speed of 5G will be 1Gbit/s.



Source: Encyclopaedia Britannica, Inc.

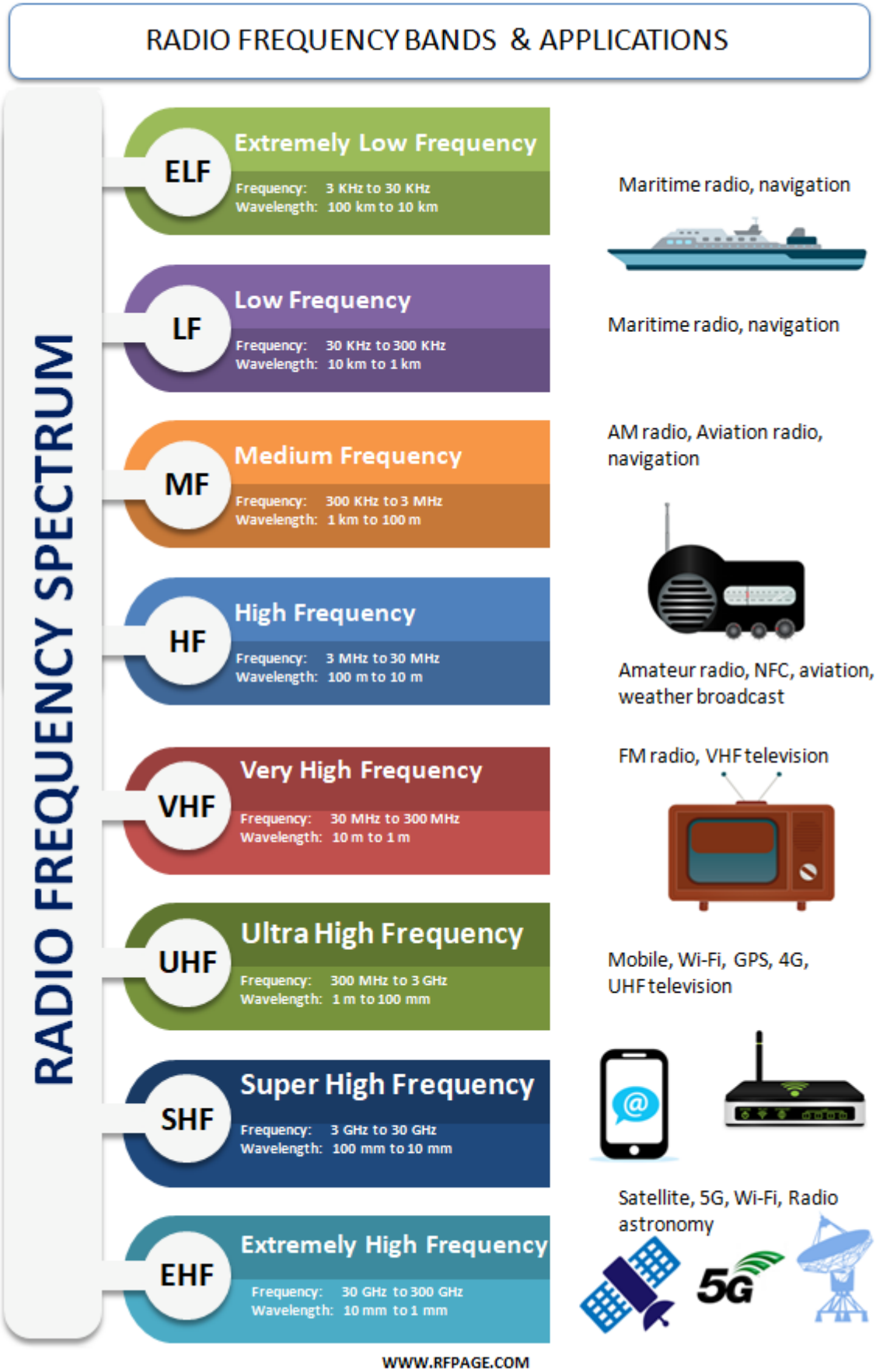
What is Radio Frequency?

RF is the lowest portion in the electromagnetic spectrum familiar as a medium of analogue and modern digital wireless communication system. It spreads in the range between 3 KHz and 300 GHz. All known transmission systems works in the RF spectrum range including analogue radio, aircraft navigation, marine radio, amateur radio, TV broadcasting, mobile networks and satellite systems. Let's take a look on each of the RF sub bands and the areas of RF spectrum uses.

Radio Frequency Spectrum: Ranges

Designation	Abbreviation	Frequencies	Wavelengths
Very Low Frequency	VLF	3 kHz - 30 kHz	100 km - 10 km
Low Frequency	LF	30 kHz - 300 kHz	10 km - 1 km
Medium Frequency	MF	300 kHz - 3 MHz	1 km - 100 m
High Frequency	HF	3 MHz - 30 MHz	100 m - 10 m
Very High Frequency	VHF	30 MHz - 300 MHz	10 m - 1 m
Ultra High Frequency	UHF	300 MHz - 3 GHz	1 m - 100 mm
Super High Frequency	SHF	3 GHz - 30 GHz	100 mm - 10 mm
Extremely High Frequency	EHF	30 GHz - 300 GHz	10 mm - 1 mm

www.rfpage.com


RADIO FREQUENCY BANDS & APPLICATIONS
RADIO FREQUENCY SPECTRUM
ELF
Extremely Low Frequency
Frequency: 3 KHz to 30 KHz
Wavelength: 100 km to 10 km
Maritime radio, navigation
LF
Low Frequency
Frequency: 30 KHz to 300 KHz
Wavelength: 10 km to 1 km
Maritime radio, navigation
MF
Medium Frequency
Frequency: 300 KHz to 3 MHz
Wavelength: 1 km to 100 m
AM radio, Aviation radio, navigation
HF
High Frequency
Frequency: 3 MHz to 30 MHz
Wavelength: 100 m to 10 m
Amateur radio, NFC, aviation, weather broadcast
VHF
Very High Frequency
Frequency: 30 MHz to 300 MHz
Wavelength: 10 m to 1 m
FM radio, VHF television
UHF
Ultra High Frequency
Frequency: 300 MHz to 3 GHz
Wavelength: 1 m to 100 mm
Mobile, Wi-Fi, GPS, 4G, UHF television
SHF
Super High Frequency
Frequency: 3 GHz to 30 GHz
Wavelength: 100 mm to 10 mm
EHF
Extremely High Frequency
Frequency: 30 GHz to 300 GHz
Wavelength: 10 mm to 1 mm
Satellite, 5G, Wi-Fi, Radio astronomy
5G
WWW.RFPAGE.COM

Signals

Signals are the physical representation of data. Users of a communication system can only exchange data through the transmission of signals. Layer 1 of the ISO/OSI basic reference model is responsible for the conversion of data, i.e., bits, into signals and vice versa (Halsall, 1996), (Stallings, 1997 and 2002). Signals are functions of time and location. Signal parameters represent the data values. The most interesting types of signals for radio transmission are periodic signals, especially sine waves as carriers.

Signal parameters are the amplitude A, the frequency f, and the phase shift φ. The amplitude as a factor of the function g may also change over time, thus At, (see section 2.6.1). The frequency f expresses the periodicity of the signal with the period T = 1/f. (In equations, ω is frequently used instead of 2πf.) The frequency f may also change over time, thus ft, (see section 2.6.2). Finally, the phase shift determines the shift of the signal relative to the same signal without a shift.

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. Somehow, we have to couple the energy from the transmitter to the outside world and, in reverse, from the outside world to the receiver. This is exactly what antennas do. 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.

However, such an antenna does not exist in reality. Real antennas all 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 λ/4 is efficient. This is also known as Marconi antenna.

Signal Propogation

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.

Path loss of radio signals In free space radio signals propagate as light does (independently of their frequency), i.e., they follow a straight line (besides gravitational effects). If such a straight line exists between a sender and a receiver it is called line-of-sight (LOS). Even if no matter exists between the sender and the receiver (i.e., if there is a vacuum), the signal still experiences the free space loss. The received power Pr is proportional to 1/d2 with d being the distance between sender and receiver (inverse square law). The reason for this phenomenon is quite simple. Think of the sender being a point in space. The sender now emits a signal with certain energy. This signal travels away from the sender at the speed of light as a wave with a spherical shape. If there is no obstacle, the sphere continuously grows with the sending energy equally distributed over the sphere's surface. This surface area s grows with the increasing distance d from the center according to the equation $s = 4\pi d2$.

Depending on the frequency, radio waves can also penetrate objects. Generally the lower the frequency, the better the penetration. Long waves can be transmitted through the oceans to a submarine while high frequencies can be blocked by a tree. The higher the frequency, the more the behavior of the radio waves resemble that of light

Radio waves can exhibit three fundamental propagation behaviors depending on their frequency:

- Ground wave (<2 MHz): Waves with low frequencies follow the earth's surface and can propagate long distances. These waves are used for, e.g., submarine communication or AM radio.
- Sky wave (2–30 MHz): Many international broadcasts and amateur radio use these short waves that are reflected2 at the ionosphere. This way the waves can bounce back and forth between the ionosphere and the earth's surface, travelling around the world.
- Line-of-sight (>30 MHz): Mobile phone systems, satellite systems, cordless telephones etc. use even higher frequencies. The emitted waves follow a (more or less) straight line of sight. This enables direct communication with satellites (no reflection at the ionosphere) or microwave links on the ground. However, an additional consideration for ground-based communication is that the waves are bent by the atmosphere due to refraction.

Multiplexing

Multiplexing is not only a fundamental mechanism in communication systems but also in everyday life. Multiplexing describes how several users can share a medium with minimum or no interference. One example, is highways with several lanes. Many users (car drivers) use the same medium (the highways) with hopefully no interference (i.e., accidents). This is possible due to the provision of several lanes (space division multiplexing) separating the traffic. In addition, different cars use the same medium (i.e., the same lane) at different points in time (time division multiplexing).

Space division multiplexing 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 si is also shown. Here space is represented via circles indicating the interference range as introduced in Figure 2.11. How is the separation of the different channels achieved? 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. In our highway example this would imply that each driver had his or her own lane. Although this procedure clearly represents a waste of space, this is exactly the principle used by the old analog telephone system: each subscriber is given a separate pair of copper wires to the local exchange. In wireless transmission, SDM implies a separate sender for each communication channel with a wide enough distance between senders. This multiplexing scheme is used, for example, at FM radio stations where the transmission range is limited to a certain region many radio stations around the world can use the same frequency without interference. Using SDM, obvious problems arise if two or more channels were established within the same space, for example, if several radio stations want to broadcast in the same city. Then, one of the following multiplexing schemes must be used (frequency, time, or code division multiplexing).

Frequency division multiplexing

Frequency division multiplexing (FDM) describes schemes to subdivide the frequency dimension into several non-overlapping frequency bands as shown in Figure 2.17. 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.

However, this scheme also has disadvantages. While radio stations broadcast 24 hours a day, mobile 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.

Time division multiplexing

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 (see Figure 2.18). 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.

Frequency and time division multiplexing can be combined, i.e., a channel ki can use a certain frequency band for a certain amount of time as shown in Figure 2.19. 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.

A disadvantage of this scheme is again the necessary coordination between different senders. One has to control the sequence of frequencies and the time of changing to another frequency. Two senders will interfere as soon as they select the same frequency at the same time. However, if the frequency change (also called frequency hopping) is fast enough, the periods of interference may be so small that, depending on the coding of data into signals, a receiver can still recover the original data.

Code division multiplexing

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, see section 2.7), it now features in many civil wireless transmission scenarios thanks to the availability of cheap processing power (explained in more detail in section 3.5). Figure 2.20 shows how all 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 technical realization of CDM is discussed in section 2.7 and chapter 3 together with the medium access mechanisms. An excellent book dealing with all aspects of CDM is Viterbi (1995). The typical everyday example of CDM is a party with many participants from different countries around the world who establish communication channels, i.e., they talk to each other, using the same frequency range (approx. 300–6000 Hz depending on a person's voice) at the same time. If everybody speaks the same language, SDM is needed to be able to communicate talking with limited transmit power). But as soon as another code, i.e., another language, is used, one can tune in to this language and clearly separate communication in this language from all the other languages. (The other languages appear as background noise.)

This explains why CDM has built-in security: if the language is unknown, the signals can still be received, but they are useless. By using a secret code (or language), a secure channel can be established in a 'hostile' environment. (At parties this may cause some confusion.) Guard spaces are also of importance in this illustrative example. Using, e.g., Swedish and Norwegian does not really work; the languages aretoo close. But Swedish and Finnish are 'orthogonal' enough to separate the communication channels. 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 (see section 3.5). 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.

Modulation

This function has three parameters: amplitude At, frequency ft, and phase φt which may be varied in accordance with data or another modulating signal. For digital modulation, which is the main topic in this section, digital data (0 and 1) is translated into an analog signal (baseband signal). Digital modulation is required if digital data has to be transmitted over a medium that only allows for analog transmission. One example for wired networks is the old analog telephone system – to connect a computer to this system a

modem is needed. The modem then performs the translation of digital data into analog signals and vice versa. Digital transmission is used, for example, in wired local area networks or within a computer (Halsall, 1996), (Stallings, 1997). In wireless networks, however, digital transmission cannot be used. Here, the binary bit-stream has to be translated into an analog signal first. The three basic methods for this translation are amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK).

Amplitude Shift Keying

Amplitude Shift Keying (ASK), the most simple digital modulation scheme. The two binary values, 1 and 0, are represented by two different amplitudes. In the example, one of the amplitudes is 0 (representing the binary 0). This simple scheme only requires low bandwidth, but is very susceptible to interference. Effects like multi-path propagation, noise, or path loss heavily influence the amplitude. In a wireless environment, a constant amplitude

Analog modulation Digital modulation Analog baseband signal Radio carrier Digital data 101101001 Figure 2.21 Modulation in a transmitter Analog demodulation Synchronization decision Analog baseband signal Radio carrier Digital data 101101001 Figure 2.22 Demodulation and data reconstruction in a receiver cannot be guaranteed, so ASK is typically not used for wireless radio transmission. However, the wired transmission scheme with the highest performance, namely optical transmission, uses ASK. Here, a light pulse may represent a 1, while the absence of light represents a 0. The carrier frequency in optical systems is some hundred THz. ASK can also be applied to wireless infra red transmission, using a directed beam or diffuse light

Frequency Shift Keying

A modulation scheme often used for wireless transmission is frequency shift keying (FSK) (see Figure 2.24). The simplest form of FSK, also called binary FSK (BFSK), assigns one frequency f1 to the binary 1 and another frequency f2 to the binary 0. A very simple way to implement FSK is to switch between two oscillators, one with the frequency f1 and the other with f2, depending on the input. To avoid sudden changes in phase, special frequency modulators with continuous phase modulation, (CPM) can be used. Sudden changes in phase cause high frequencies, which is an undesired side-effect.

A simple way to implement demodulation is by using two bandpass filters, one for f1 the other for f2. A comparator can then compare the signal levels of the filter outputs to decide which of them is stronger. FSK needs a larger bandwidth compared to ASK but is much less susceptible to errors.

Phase Shift Keying

Finally, phase shift keying (PSK) uses shifts in the phase of a signal to represent data. Figure 2.25 shows a phase shift of 180° or π as the 0 follows the 1 (the same happens as the 1 follows the 0). This simple scheme, shifting the phase by 180° each time the value of data changes, is also called binary PSK (BPSK). A simple implementation of a BPSK modulator could multiply a frequency f with +1 if the binary data is 1 and with –1 if the binary data is 0. To receive the signal correctly, the receiver must synchronize in frequency and phase with the transmitter. This can be done using a phase lock loop (PLL). Compared to FSK, PSK is more resistant to interference, but receiver and transmitter are also more complex.

Spread Spectrum

Spread spectrum techniques involve spreading the bandwidth needed to transmit data – which does not make sense at first sight. Spreading the bandwidth has several advantages. The main advantage is the resistance to narrowband interference. In Figure 2.32, diagram

1. shows an idealized narrowband signal from a sender of user data (here power density dP/df versus frequency f). The sender now spreads the signal in step ii)
2. converts the narrowband signal into a broadband signal. The energy needed to transmit the signal (the area shown in the diagram) is the same, but it is now spread over a larger frequency range. The power level of the spread signal can be much lower than that of the original narrowband signal without losing data. Depending on the generation and reception of the spread signal, the power level of the user signal can even be as low as the background noise. This makes it difficult to distinguish the user signal from the background noise and thus hard to detect. During transmission, narrowband and broadband interference add to the signal in step iii).
3. The sum of interference and user signal is received. The receiver now knows how to despread the signal, converting the spread user signal into a narrowband signal again, while spreading the narrowband interference and leaving the broadband interference.
4. the receiver applies abandpass filter to cut off frequencies left and right of the narrowband signal.Finally, the receiver can reconstruct the original data because the power level of the user signal is high enough, i.e., the signal is much stronger than the remaining interference.

Cellular System

Cellular systems for mobile communications implement SDM. Each transmitter, typically called a base station, covers a certain area, a cell. Cell radii can vary from tens of meters in buildings, and hundreds of meters in cities, up to tens of kilometers in the countryside. The shape of cells are never perfect circles or hexagons (as shown in Figure 2.41), but depend on the environment (buildings, mountains, valleys etc.), on weather conditions, and sometimes even on system load. Typical systems using this approach are mobile telecommunication systems (see chapter 4), where a mobile station within the cell around a base station communicates with this base station and vice versa.

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

Small cells also have some disadvantages:

- Infrastructure needed: Cellular systems need a complex infrastructure to connect all base stations. This includes many antennas, switches for call forwarding, location registers to find a mobile station etc, which makes the whole system quite expensive.

- Handover needed: The mobile station has to perform a handover when changing from one cell to another. Depending on the cell size and the speed of movement, this can happen quite often.
- Frequency planning: To avoid interference between transmitters using the same frequencies, frequencies have to be distributed carefully. On the one hand, interference should be avoided, on the other, only a limited number of frequencies is available.

Telecommunication, Satellite and Broadcast Systems: GSM

GSM is the most successful digital mobile telecommunication system in the world today. It is used by over 800 million people in more than 190 countries. In the early 1980s, Europe had numerous coexisting analog mobile phone systems, which were often based on similar standards (e.g., NMT 450), but ran on slightly different carrier frequencies. To avoid this situation for a second generation fully digital system, the groupe spéciale mobile (GSM) was founded in 1982. This system was soon named the global system for mobile communications (GSM), with the specification process lying in the hands of ETSI (ETSI, 2002), (GSM Association, 2002). In the context of UMTS and the creation of 3GPP (Third generation partnership project, 3GPP, 2002a) the whole development process of GSM was transferred to 3GPP and further development is combined with 3G development. 3GPP assigned new numbers to all GSM standards. However, to remain consistent with most of the GSM literature, this GSM section stays with the original numbering (see 3GPP, 2002a, for conversion).

The primary goal of GSM was to provide a mobile phone system that allows users to roam throughout Europe and provides voice services compatible to ISDN and other PSTN systems. The specification for the initial system already covers more than 5,000 pages; new services, in particular data services, now add even more specification details. Readers familiar with the ISDN reference model will recognize many similar acronyms, reference points, and interfaces. GSM standardization aims at adopting as much as possible.

Mobile Services

GSM permits the integration of different voice and data services and the interworking with existing networks. Services make a network interesting for customers. GSM has defined three different categories of services: bearer, tele, and supplementary services. These are described in the following subsections. Figure 4.3 shows a reference model for GSM services. A mobile station MS is connected to the GSM public land mobile network (PLMN) via the Um interface. (GSM-PLMN is the infrastructure needed for the GSM network.) This network is connected to transit networks, e.g., integrated services digital network (ISDN) or traditional public switched telephone network (PSTN). There might be an additional network, the source/destination network, before another terminal TE is connected. Bearer services now comprise all services that enable the transparent transmission of data between the interfaces to the network, i.e., S in case of the mobile station, and a similar interface for the other terminal (e.g., S0 for ISDN terminals). Interfaces like U, S, and R in case of ISDN have not been defined for all networks, so it depends on the specific network which interface is used as a reference for the transparent transmission of data. In the classical GSM model, bearer services are connection-oriented and circuit- or packet-switched. These services only need the lower three layers of the ISO/OSI reference model.

Within the mobile station MS, the mobile termination (MT) performs all network specific tasks (TDMA, FDMA, coding etc.) and offers an interface for data transmission (S) to the terminal TE which can then be network independent. Depending on the capabilities of TE, further interfaces may be needed, such as R, according to the ISDN reference model (Halsall, 1996). Tele services are application specific and may thus need all seven layers of the ISO/OSI reference model. These services are specified end-to-end, i.e., from one terminal TE to another.

1. **Bearer services** GSM specifies different mechanisms for data transmission, the original GSM allowing for data rates of up to 9600 bit/s for non-voice services. Bearer services permit transparent and non-transparent, synchronous or asynchronous data transmission. Transparent bearer services only use the functions of the physical layer (layer 1) to transmit data. Data transmission has a constant delay and throughput if no transmission errors occur. The only mechanism to increase transmission quality is the use of forward error correction (FEC), which codes redundancy into the data stream and helps to reconstruct the original data in case of transmission errors. Depending on the FEC, data rates of 2.4, 4.8, or 9.6 kbit/s are possible. Transparent bearer services do not try to recover lost data in case of, for example, shadowing or interruptions due to handover. Non-transparent bearer services use protocols of layers two and three to implement error correction and flow control. These services use the transparent bearer services, adding a radio link protocol (RLP). This protocol comprises mechanisms of high-level data link control (HDLC), (Halsall, 1996) and special selective-reject mechanisms to trigger retransmission of erroneous data. The achieved bit error rate is less than 10–7, but now throughput and delay may vary depending on transmission quality.
2. **Tele services** GSM mainly focuses on voice-oriented tele services. These comprise encrypted voice transmission, message services, and basic data communication with terminals as known from the PSTN or ISDN (e.g., fax). However, as the main service is telephony, the primary goal of GSM was the provision of high-quality digital voice transmission, offering at least the typical bandwidth of 3.1 kHz of analog phone systems. Special codecs (coder/decoder) are used for voice transmission, while other codecs are used for the transmission of analog data for communication with traditional computer modems used in, e.g., fax machines. Another service offered by GSM is the emergency number. The same number can be used throughout Europe. This service is mandatory for all providers and free of charge. This connection also has the highest priority, possibly pre-empting other connections, and will automatically be set up with the closest emergency center. A useful service for very simple message transfer is the short message service (SMS), which offers transmission of messages of up to 160 characters. SMS messages do not use the standard data channels of GSM but exploit unused capacity in the signalling channels (see section 4.1.3.1). Sending and receiving of SMS is possible during data or voice transmission. SMS was in the GSM standard from the beginning; however, almost no one used it until millions of young people discovered this service in the mid-nineties as a fun service.
3. **Supplementary services** In addition to tele and bearer services, GSM providers can offer supplementary services. Similar to ISDN networks, these services offer various enhancements for the standard telephony service, and may vary from provider to provider. Typical services are user identification, call redirection, or forwarding of ongoing calls. Standard ISDN features such as closed user groups and multiparty communication may be available. Closed user groups are of special interest to companies because they allow, for example, a company-specific GSM sub-network, to which only members of the group have access.

Architecture

GSM comes with a hierarchical, complex system architecture comprising many entities, interfaces, and acronyms. Figure 4.4 gives a simplified overview of the GSM system as specified in ETSI (1991b). A GSM system consists of three subsystems, the radio sub system (RSS), the network and switching subsystem (NSS), and the operation subsystem (OSS). Each subsystem will be discussed in more detail in the following sections. Generally, a GSM customer only notices a very small fraction of the whole network – the mobile stations (MS) and some antenna masts of the base transceiver stations (BTS).

1. **Radio subsystem** As the name implies, the radio subsystem (RSS) comprises all radio specific entities, i.e., the mobile stations (MS) and the base station subsystem (BSS). Figure 4.4 shows the connection between the RSS and the NSS via the A interface (solid lines) and the connection to the OSS via the O

interface (dashed lines). The A interface is typically based on circuit-switched PCM-30 systems (2.048 Mbit/s), carrying up to 30 64 kbit/s connections, whereas the O interface uses the Signalling System No. 7 (SS7) based on X.25 carrying management data to/from the RSS.

- Base station subsystem (BSS): A GSM network comprises many BSSs, each controlled by a base station controller (BSC). The BSS performs all functions necessary to maintain radio connections to an MS, coding/decoding of voice, and rate adaptation to/from the wireless network part. Besides a BSC, the BSS contains several BTSs.

- Base transceiver station (BTS): A BTS comprises all radio equipment, i.e., antennas, signal processing, amplifiers necessary for radio transmission. A BTS can form a radio cell or, using sectorized antennas, several cells (see section 2.8), and is connected to MS via the Um interface (ISDN U interface for mobile use), and to the BSC via the Abis interface. The Um interface contains all the mechanisms necessary for wireless transmission (TDMA, FDMA etc.) and will be discussed in more detail below. The Abis interface consists of 16 or 64 kbit/s connections. A GSM cell can measure between some 100 m and 35 km depending on the environment (buildings, open space, mountains etc.) but also expected traffic.

- Base station controller (BSC): The BSC basically manages the BTSs. It reserves radio frequencies, handles the handover from one BTS to another within the BSS, and performs paging of the MS. The BSC also multiplexes the radio channels onto the fixed network connections at the A interface.

- Mobile station (MS): The MS comprises all user equipment and software needed for communication with a GSM network. An MS consists of user independent hard- and software and of the subscriber identity module (SIM), which stores all user-specific data that is relevant to GSM.3 While an MS can be identified via the international mobile equipment identity (IMEI), a user can personalize any MS using his or her SIM, i.e., user-specific mechanisms like charging and authentication are based on the SIM, not on the device itself. Device-specific mechanisms, e.g., theft protection, use the device specific IMEI. Without the SIM, only emergency calls are possible. The SIM card contains many identifiers and tables, such as card-type, serial number, a list of subscribed services, a personal identity number (PIN), a PIN unblocking key (PUK), an authentication key Ki, and the international mobile subscriber identity (IMSI) (ETSI, 1991c). The PIN is used to unlock the MS. Using the wrong PIN three times will lock the SIM. In such cases, the PUK is needed to unlock the SIM. The MS stores dynamic information while logged onto the GSM system, such as, e.g., the cipher key Kc and the location information consisting of a temporary mobile subscriber identity (TMSI) and the location area identification (LAI). Typical MSs for GS 900 have a transmit power of up to 2 W, whereas for GSM 1800 1 W is enough due to the smaller cell size.

2. **Network and Switching Subsystem (NSS)**

The “heart” of the GSM system is formed by the network and switching subsystem (NSS). The NSS connects the wireless network with standard public networks, performs handovers between different BSSs, comprises functions for worldwide localization of users and supports charging, accounting, and roaming of users between different providers in different countries. The NSS consists of the following switches and databases:

- **Mobile services switching center (MSC):** MSCs are high-performance digital ISDN switches. They set up connections to other MSCs and to the BSCs via the A interface, and form the fixed backbone network of a GSM system. Typically, an MSC manages several BSCs in a geographical region. A gateway MSC (GMSC) has additional connections to other fixed networks, such as PSTN and ISDN. Using additional interworking functions (IWF), an MSC can also connect to public data networks (PDN) such as X.25. An MSC handles all signaling needed for connection setup, connection release and handover of connections to other MSCs. The standard signaling system No. 7 (SS7) is used for this purpose. SS7 covers all aspects of control signalingfor digital networks (reliable routing and delivery of control messages, establishing

and monitoring of calls). Features of SS7 are number portability, free phone/toll/collect/credit calls, call forwarding, three-way calling etc. An MSC also performs all functions needed for supplementary services such as call forwarding, multi-party calls, reverse charging etc.

- **Home location register (HLR):** The HLR is the most important database in a GSM system as it stores all user-relevant information. This comprises static information, such as the mobile subscriber ISDN number (MSISDN), subscribed services (e.g., call forwarding, roaming restrictions, GPRS), and the international mobile subscriber identity (IMSI). Dynamic information is also needed, e.g., the current location area (LA) of the MS, the mobile subscriber roaming number (MSRN), the current VLR and MSC. As soon as an MS leaves its current LA, the information in the HLR is updated. This information is necessary to localize a user in the worldwide GSM network. All these user-specific information elements only exist once for each user in a single HLR, which also supports charging and accounting. The parameters will be explained in more detail in section 4.1.5. HLRs can manage data for several million customers and contain highly specialized data bases which must fulfill certain real-time requirements to answer requests within certain time-bounds.
- **Visitor location register (VLR):** The VLR associated to each MSC is a dynamic database which stores all important information needed for the MS users currently in the LA that is associated to the MSC (e.g., IMSI, MSISDN, HLR address). If a new MS comes into an LA the VLR is responsible for, it copies all relevant information for this user from the HLR. This hierarchy of VLR and HLR avoids frequent HLR updates and long-distance signaling of user information. The typical use of HLR and VLR for user localization will be described in section 4.1.5. Some VLRs in existence, are capable of managing up to one million customers.

3. **Operation subsystem** The third part of a GSM system, the operation subsystem (OSS), contains the necessary functions for network operation and maintenance. The OSS possesses network entities of its own and accesses other entities via SS7 signaling (see Figure 4.4). The following entities have been defined:

- **Operation and maintenance center (OMC):** The OMC monitors and controls all other network entities via the O interface (SS7 with X.25). Typical OMC management functions are traffic monitoring, status reports of network entities, subscriber and security management, or accounting and billing. OMCs use the concept of telecommunication management network (TMN) as standardized by the ITU-T.
- **Authentication centre (AuC):** As the radio interface and mobile stations are particularly vulnerable, a separate AuC has been defined to protect user identity and data transmission. The AuC contains the algorithms for authentication as well as the keys for encryption and generates the values needed for user authentication in the HLR. The AuC may, in fact, be situated in a special protected part of the HLR.
- **Equipment identity register (EIR):** The EIR is a database for all IMEIs, i.e., it stores all device identifications registered for this network. As MSs are mobile, they can be easily stolen. With a valid SIM, anyone could use the stolen MS. The EIR has a blacklist of stolen (or locked) devices. In theory an MS is useless as soon as the owner has reported a theft. Unfortunately, the blacklists of different providers are not usually synchronized and the illegal use of a device in another operator's network is possible (the reader may speculate as to why this is the case). The EIR also contains a list of valid IMEIs (white list), and a list of malfunctioning devices (gray list).

Radio Interface

The most interesting interface in a GSM system is Um, the radio interface, as it comprises many mechanisms for multiplexing and media access. GSM implements SDMA using cells with BTS and assigns an MS to a BTS. Furthermore, FDD is used to separate downlink and uplink.Media access combines TDMA and FDMA. In GSM 900, 124 channels, each 200 kHz wide, are used for FDMA, whereas GSM 1800 uses, 374 channels. Due to technical reasons, channels 1 and 124 are not used for transmission in GSM 900. Typically, 32 channels are reserved for organizational data; the remaining 90 are used for customers. Each BTS then manages a single

channel for organizational data and, e.g., up to 10 channels for user data. The following example is based on the GSM 900 system, but GSM works in a similar way at 1800 and 1900 MHz.

Each of the 248 channels is additionally separated in time via a GSM TDMA frame, i.e., each 200 kHz carrier is subdivided into frames that are repeated continuously. The duration of a frame is 4.615 ms. A frame is again subdivided into 8 GSM time slots, where each slot represents a physical TDM channel and lasts for 577 µs. Each TDM channel occupies the 200 kHz carrier for 577 µs every 4.615 ms.

Data is transmitted in small portions, called bursts.the burst is only 546.5 µs long and contains 148 bits. The remaining 30.5 µs are used as guard space to avoid overlapping with other bursts due to different path delays and to give the transmitter time to turn on and off. Filling the whole slot with data allows for the transmission of156.25 bit within 577 µs. Each physical TDM channel has a raw data rate of about 33.8 kbit/s, each radio carrier transmits approximately 270 kbit/s over the Um interface.

The first and last three bits of a normal burst (tail) are all set to 0 and can be used to enhance the receiver performance. The training sequence in the middle of a slot is used to adapt the parameters of the receiver to the current path propagation characteristics and to select the strongest signal in case of multi-path propagation. A flag S indicates whether the data field contains user or network control data. Apart from the normal burst, ETSI (1993a) defines four more bursts for data transmission: a frequency correction burst allows the MS to correct the local oscillator to avoid interference with neighboring channels, a synchronization burst with an extended training sequence synchronizes the MS with the BTS in time, an access burst is used for the initial connection setup between MS and BTS, and finally a dummy burst is used if no data is available for a slot.

Two factors allow for the use of simple transmitter hardware: on the one hand, the slots for uplink and downlink of a physical TDM channel are separated in frequency (45 MHz for GSM 900, 95 MHz for GSM 1800 using FDD). On the other hand, the TDMA frames are shifted in time for three slots, i.e., if the BTS sends data at time t0 in slot one on the downlink, the MS accesses slot one on the uplink at time t0+3·577 µs. An MS does not need a full-duplex transmitter, a simpler half-duplex transmitter switching between receiving and sending is enough. To avoid frequency selective fading, GSM specifies an optional slow frequency hopping mechanism. MS and BTS may change the carrier frequency after each frame based on a common hopping sequence. An MS changes its frequency between up and downlink slots respectively.