Cellular Networks: Past, Present and Future

LO Walters and PS Kritzinger

March 14, 2000

1 Abstract

Mobile communication is the fastest growing field in the telecommunications industry. The most successful mobile communication system is the cellular radio network. A cellular radio network is used to transmit voice and data. Data transmission over a cellular network is a new service, which makes data networks accessible from mobile terminals via cellular telephones. This article discusses cellular radio network technology and its applications, especially pertaining to data transmission. Its history, present state, operation and planned future developments are explained.

2 Introduction

This article is organised as follows: Section 3 puts cellular radio networks in context with other mobile communication technologies. A short description of each of the most widely used mobile communication technologies is given. Section 4 tells the history of cellular radio networks. Section 5 reports on the current state of cellular technology. Section 6 explains how they work. Section 7 gives an overview of the protocols used. Section 8 explains how the radio link works. Section 9 outlines the exciting new development of high speed data transmission over cellular radio networks. Section 10 gives an overview of value added services.

3 Existing Mobile Communication Technologies

Mobility in communication is achieved by means of wireless network technology. Wireless communication is achieved by means of using radio waves. The most successful mobile radio communication systems are the following:

Paging A simple and inexpensive form of mobile communication. An antenna or satellite broadcasts short messages to subscribers. Receivers are usually devices such as beepers which display messages on a small screen. Transmission of data is one-way [Tan96]. Paging systems are designed to provide reliable communication to subscribers wherever they are. This necessitates large transmitter powers and low data rates for maximum coverage of each transmitter's designated area [Rap96].

Communication Satellites Satellites consist out of large transponders which listen to a particular radio frequency, amplify the signal, and then rebroadcast it at another frequency. They are therefore inherently broadcast devices. A drawback of satellites is that they have quite a large propagation delay due to the distances travelled by radio waves [Tan96].

Cellular Radio Networks Communication in a cellular network is two-way. Cellular phones communicate with tranceivers by sending and receiving messages on two different frequencies. A geographic area is divided up into cells, each cell being serviced by a tranceiver or base station. Adjacent cells use different frequencies to communicate with phones. Cells that are one cell apart can use the same frequencies. In this way adjacent cells don't interfere with one another, and frequencies can be reused. Radio signals within a cell can be transmitted strongly, as they won't interfere with adjacent cells. This ensures that quality of transmission is good[Tan96].

Personal Handyphone The PHS - Personal Handyphone System is used in Japan. It is similar to cellular networks, but phones can also communicate directly with one another when in range. This is an advantage over cellular phones, which can only communicate with one another via base station tranceivers. This system is very popular within heavily populated metropolitan areas [GSM].

Mobile Radio Mobile radio is in many ways the predecessor of the cellular radio network. It is mostly analogue, and makes use of single frequencies for sending and receiving signals. Communication is half-duplex, and a button has to be pressed to switch mode. It is mostly used by applications for emergency services, the transport sector and security industry [GSM].

4 History of Cellular Radio Networks

In 1946, the first car-based telephone was set up in St. Louis in the USA. The system used a single radio transmitter on top of a tall building. A single channel was used, and therefore a button was pushed to talk, and released to listen [Tan96]. This half duplex system is still used by modern day CB-radio systems used by police and taxi operators. In the 60's the system was improved to a two-channel system, called improved mobile telephone system (IMTS)[Tan96]. The system could not support many users as frequencies were limited.

This problem was solved by the idea of using cells to facilitate the reuse of frequencies. More users can be supported in such a cellular radio system. It was implemented for the first time in the advanced mobile phone system (AMPS). AMPS was and still is an analogue system, and is part of first generation cellular radio systems. Second generation systems are digital. In the USA two standards are used for second generation systems - IS-95 (CDMA) and IS-136 (D-AMPS) [Tan96] [GSM]. Europe consolidated on one system called global system for mobile communications (GSM)[GSM]. Japan uses a system called personal digital cellular (PDC).

5 Present Situation

Cellular radio is the fastest growing segment of the communications industry today [AMH+99]. Cellular companies reported a subscription base of more than 200 million people in 1997. This figure reportedly grows by an average of 150 000 new subscribers every day [AMH+99].

Because of Europe's early commitment to one system, it is leading the field in both its subscriber base and data transmission capabilities. GSM is used in over 100 countries by over 215 operators inside and outside of Europe [GSM]. The Japanese PDC system is the second largest digital cellular system after GSM. Table 1 compares the air interface characteristics of the four largest digital cellular radio systems in use today.

Current cellular radio systems are in their second generation (2G). The third generation of cellular systems (3G systems) will allow different systems to interoperate in order to attain global

	GSM	IS-54/136	IS-95	PDC
Region	Europe	North America	North America	Japan
Frequency band (MHz)	900/1800/1900	800/1900	800/1900	700/1500
Multiple access	F/TDMA	F/TDMA	F/CDMA	F/TDMA

Table 1: Air interfaces of the largest digital cellular radio systems in use

roaming across different cellular radio networks [ZAB99]. The International Telecommunication Union (ITU) has been doing research on 3G systems since the mid 80's. Their version of a 3G system is called international mobile telecommunications - 2000 (IMT-2000).

European countries are doing research into 3G systems under the auspices of the European Community [ZAB99]. Their system is referred to as the universal mobile telecommunication system (UMTS). It has the same goals as the IMT-2000 system. The major objectives of 3G systems are:

- Use of common global frequencies for all cellular networks.
- Worldwide roaming.
- Standardization of radio interfaces.
- High data transmission rates for both circuit and packet switched data.
- Efficient spectrum utilization schemes.

6 How Does it Work?

In the explanation that follows a cellular telephone or any other device that can connect to a cellular radio network will be referred to as a mobile station. This is in keeping with the literature on the subject.

The reason for the cellular topology of a cellular radio network is to facilitate the efficient use of limited radio spectrum. The cellular topology allows the same frequencies to be reused in different cells.

As mentioned before, a geographic region is divided up into cells. Each cell has a base station which transmits data via a radio link to mobile stations within the cell. Figure 1 illustrates how this works. The base stations are connected to the mobile switching center (MSC) via microwave links or telephone lines. The MSC connects to the public switched telephone network (PSTN), which routes calls to other mobile stations or land based telephones.

A cellular network therefore makes use of both land and radio based network links. Because of this, it is often referred to as a PLMN - Public Land Mobile Network [AMH+99].

The operation of a cellular network system can be explained by an example of a mobile station placing a call to another mobile station:

A mobile station places a call by sending a call initiation request to its nearest base station. This request is sent on a special channel called the reverse control channel (RCC). The base station sends this request on to the MSC, which validates the request. Contained in the request is the telephone number of the called party. The MSC uses this number to make a connection to the called party via the PSTN. It first connects itself to the MSC of the called party. It then instructs the base station and mobile station which placed the call to switch to voice channels.

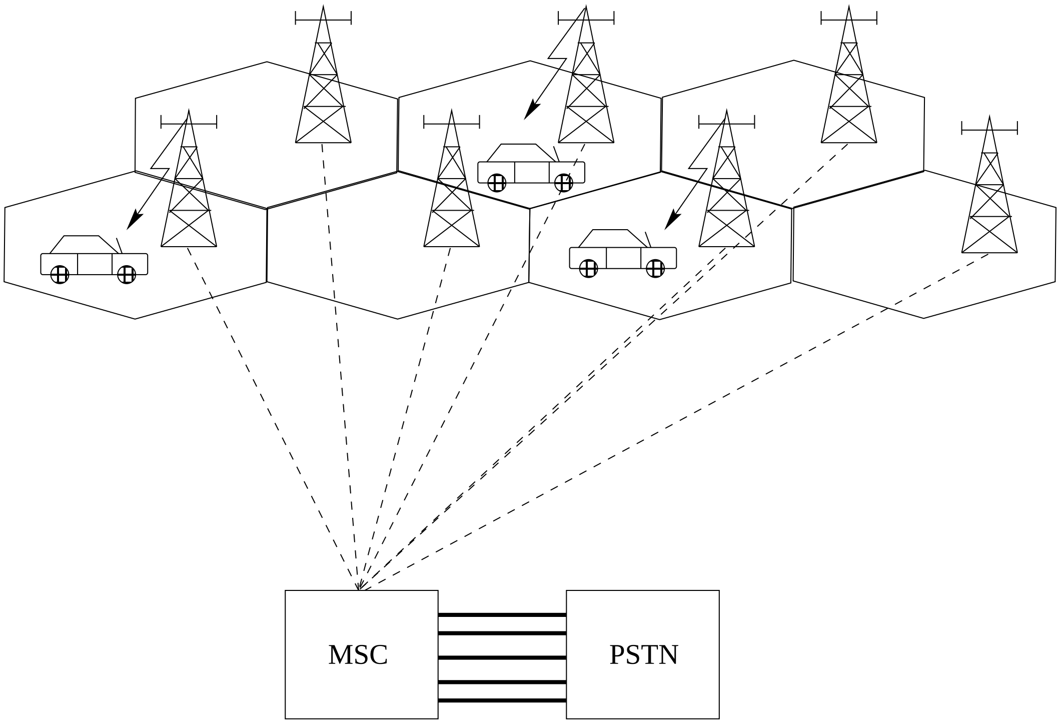

Figure 1: A cellular system

The mobile station which placed the call is then connected to the called station on unused forward and backward voice channels (FVC, BVC) [Rap96].

The steps that take place when a mobile station receives an incoming call are as follows:

Mobile stations continually scan the forward control channel (FCC) for paging signals from base stations. When a MSC receives a request for connection to a mobile station in its area, it sends a broadcast message to all base stations under its control. The message contains the number of the mobile station which is being called. The base stations then broadcasts this message on all forward control channels (FCC's). The correct mobile station acknowledges this page, by identifying itself over the reverse control channel (RCC). The MSC receives this acknowledgement via the base station, and instructs the base station and mobile station to switch to an unused voice channel. A data message is then transmitted over the forward voice channel, which instructs the mobile phone to ring [Rap96].

All of the steps explained above happen fast enough for the user not to experience any noticeable delay between placing a request for a call and the call being connected.

7 The Protocols Used

A communication protocol is a set of messages and rules according to which two or more entities exchange information. A cellular radio network is made up of different entities which communicate with each other i.e. the mobile station, the base station, the mobile switching centre and others which are not covered in this article. The protocols which are used by the network are organised into a protocol stack.

The GSM protocol stack is shown in Figure 2. The three entities communicating are the mobile station (MS), the base station subsystem (BSS) and the mobile switching centre (MSC).

They communicate with one another over the radio interface and the A-interface respectively.

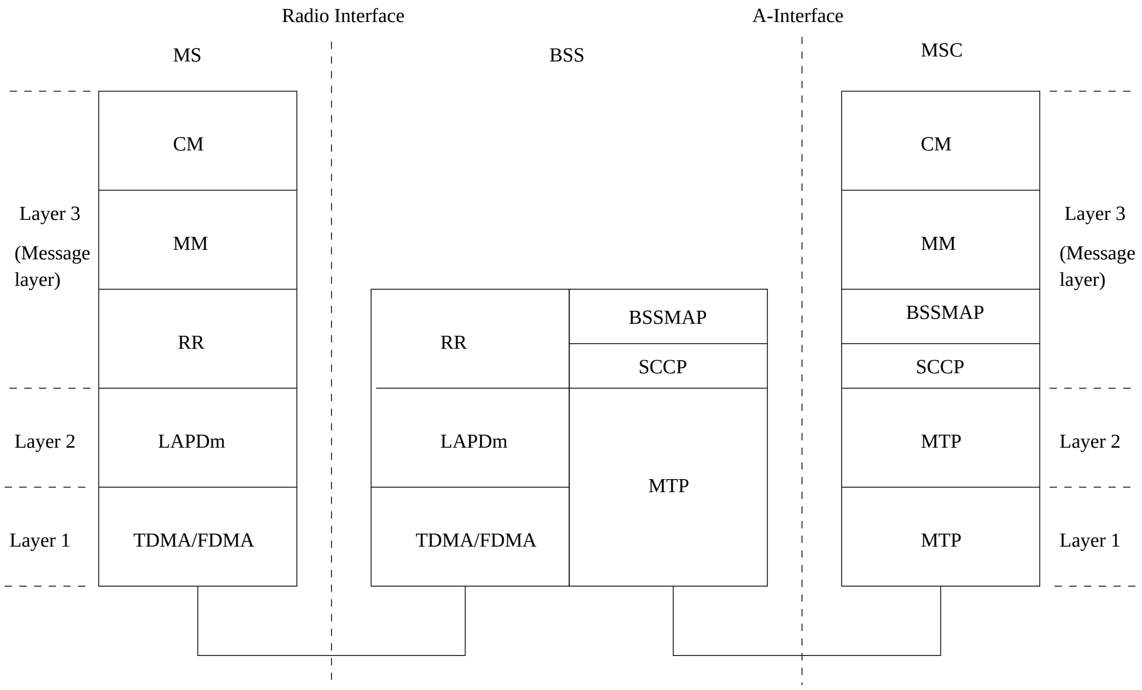

Figure 2: GSM protocol stack

The protocol stack is composed of 3 layers. Layers 1 and 2 are the physical and link layers respectively. Layer 3 is composed of 3 parts. It is known as the message or signaling layer. The physical and link layers will merely be introduced in this section. The message layer is discussed in more detail.

The physical layer specifies how data is transmitted from one entity to another over the physical transport medium. The transport layer between the MS and BSS is the radio link. The transport medium between the BSS and the MSC is either a landline or microwave based E1 trunk [Rah93].

The physical layer between the MS and the BSS is discussed in Section 8. The physical layer between the BSS and the MSC is the MTP level 1 protocol of the SS7 protocol suite [Rah93].

The link layer provides error detection and correction. An LAPD-like protocol is used over the radio interface. The MTP level 2 protocol of the SS7 protocol suite is used over the A-interface.

The three different parts of the message layer manages the resources of a cellular radio network. The parts are:

1. Radio resource management (RR)
2. Mobility management (MM)
3. Call management (CM)

The next three sections discuss each of these parts in more detail.

7.1 Radio Resource Management

The radio link between the MS and the BSS is the defining element of mobile wireless communication systems. The radio link makes it possible for users to be mobile whilst communicating.

The effective use of the radio spectrum is the primary issue here. The cellular configuration of the network is designed as such to make effective use of limited radio resources. By using different frequencies in adjacent cells, frequencies can be reused in cells further apart. This allows a maximum number of communication channels to be created by using a fixed amount of radio bandwidth.

The radio resource management layer controls the radio link between the MS and the BSS. The setup of calls as well as the maintenance of them is controlled.

A control channel exists permanently between mobile stations and base stations. When a request for call setup is made, the radio resource control layer sets up a dedicated voice circuit.

Once a voice circuit exists, it needs to be managed. Mobile stations move from one cell to another. When this happens while a call is in progress, the voice circuit needs to be transferred from one base station to another. The MSC and mobile station use the radio resource management layer to coordinate this procedure. The procedure is called handover of handoff, and is described in [Sco97], [Rap96] and [AMH+99].

7.2 Mobility Management

A cellular radio network consists of clusters of cells, each controlled by a MSC. Figure 3 shows this configuration. It is necessary for the network to know where every registered mobile station is within this configuration in order to connect it on request. The management of mobile station location information is handled by the mobility management scheme. Any registered mobile station's current BSS can be located by means of this scheme. The scheme makes use of a distributed database, and an associated protocol to maintain the location information of all the mobile stations. The protocol is the MM part of layer 3.

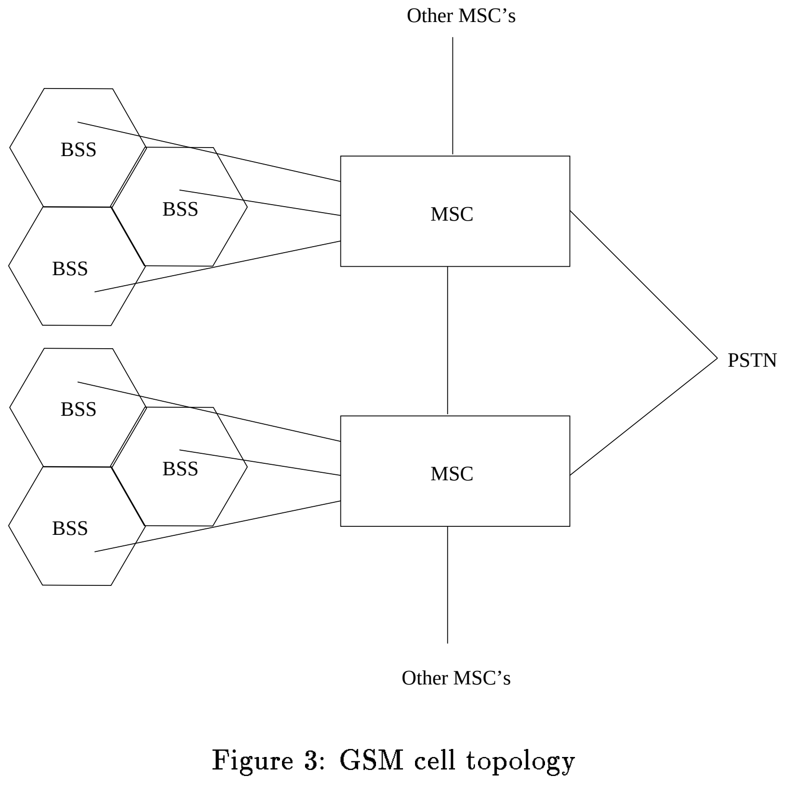

Figure 3: GSM cell topology

The distributed database consists of two different parts, a visiting location register (VLR) and a home location register (HLR). A VLR is kept at each MSC, and one HLR is kept centrally at the service provider's premises. The VLR stores a list of all the mobile stations currently in a MSC's area. The HLR has an entry for each mobile station registered with the service provider. Each mobile station's entry in the HLR is associated with an entry which gives the name of the MSC the mobile station is currently registered with.

Using this scheme, a query to the centralized HLR is all that is necessary to find the current location of a mobile station. In order to keep the HLR current, a considerable amount of information is transmitted over the network. The HLR is updated every time a mobile station moves from one MSC's area into another. As the number of subscribers on the network increases, the information flow because of updates to the HLR increases exponentially. Refer to [AMH$^+$99] for information on mobility management techniques which minimize information flow by means of optimized algorithms.

Mobility management also concerns itself with the authentication of mobile stations to the network. Mobile stations entering an MSC's area must be authenticated before they can make use of the network's resources. This ensures that only valid customers make use of the network [AMH$^+$99] [Sco97].

7.3 Communication Management

The communication management part of layer 3 is concerned with the setting up of calls. A large part of the call setup procedure is concerned with accommodating roaming on cellular network systems. Roaming is a facility that allows users to call one another from countries which have service provider's other than their own. The foreign service provider contacts a Gateway MSC (GMSC), which links it to the user's own service provider.

The GMSC identifies a caller's service provider, by looking it up in a table. It makes use of a mobile station ISDN number (MISDN), which uniquely identifies a mobile station [Sco97]. The MSISDN number is shown in Figure 4. It consists of a country code (CC), national destination code (NDC) and a subscriber number (SN). After having identified the caller's service provider, the service provider's HLR is queried to establish the location of the called mobile station. The CM part of layer 3 provides the protocol for doing all this.

CC	NDC	SN

Figure 4: The structure of the GSM MSISDN

8 Radio Link Overview

The physical layer between the mobile station and the base station uses a radio link. Information is transmitted digitally on this link. In order to transmit speech, it is encoded into a digital form. The algorithm which is used to do the encoding is called the *regular pulse excited, linear predictive coder* encoding scheme (RPE-LPC) [Sco97].

The reason why this algorithm is used rather than the well known pulse code modulation (PCM) scheme, is that it is more efficient than PCM. To transmit voice in PCM format, a link with transmission rate of 64Kbps is needed. Current GSM networks have a maximum data transmission rate of 9.6 Kbps.

Bandwidth on the radio link is constrained by limited available radio spectrum. The ITU World Radio Conference divided up the available radio spectrum between users [AMH+99] in 1995. Cellular networks were allocated approximately 25MHz of radio spectrum by this Conference [AMH+99].

A cellular radio network uses its allocated spectrum by dividing it up into 124 carrier frequencies. This is done by making use of frequency division multiple access (FDMA). These carrier frequencies are allocated to the different base stations. A base station divides each of its carrier frequencies up into eight communication channels using time division multiple access (TDMA). TDMA makes use of the fact that digital data is transmitted in frames across the radio link. Each radio link frame of specified length is broken up into eight smaller TDMA frames. Each one of the smaller frames form a communication channel. Figure 5 shows how one carrier frequency is divided up into eight communication channels by means of TDMA. Only one channel out of the 8 is used for voice traffic. The other channels are used as follows:

1. Broadcast Control - (BCCH)
2. Frequency Correction - (FCCH)
3. Synchronisation - (SCH)
4. Random Accessing - (RACH)
5. Paging - (PCH)
6. Access Granting - (AGCH)

One channel is currently unused. A description of each of the abovementioned channels can be found in Scourias [Sco97]. Using this structure a single carrier frequency can be used to transmit voice and data at 9,6 kbps.

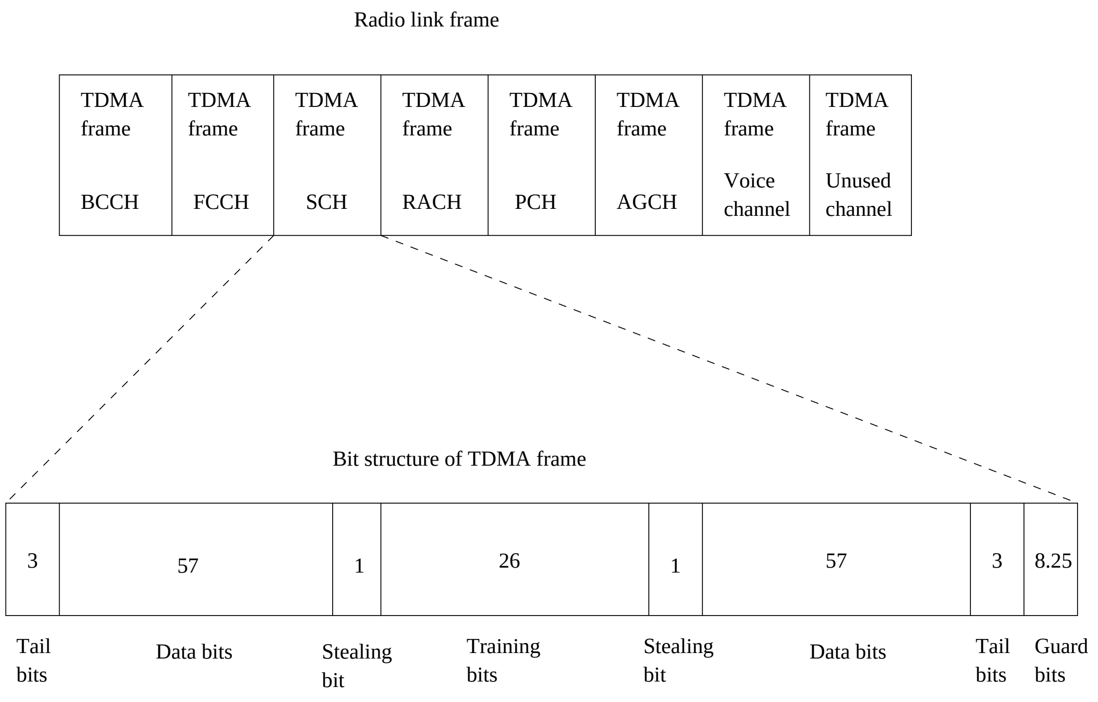

Figure 5: TDMA division of a single bearer channel into eight communication channels

9 The Future

"The holy grail of the telephone world is a small cordless phone that you can use around the house and take with you anywhere in the world - it should respond to the same number, no matter where you are", is what Tanenbaum [Tan96] defined PCS (Personal Communication Services) to be in 1996. Since then the meaning of the term has evolved. It now takes into consideration the fact that data services have changed the landscape of mobile communications. Current GSM networks not only provide voice services, but data services at 9,6 kbps. The meaning of PCS now not only encompasses the delivery of voice services to people regardless of where they are but also that of data services regardless of location, network or terminal used [HC98].

The universal mobile telecommunications system (UMTS) proposes to achieve this ideal of PCS data services that are independent of location, network or terminal. This system is hailed as being the next generation cellular system, and is referred to as 3G Cellular [Rys99a]. In order to attain this ideal, all interest holders of the cellular world will have to work together to standardize communications.

The ideal is a tall order. Other than integrating telecommunications networks, it envisions data transfer rates of up to 2Mbps [HC98]. Although the ideal won't be realised in the next few years, cellular networks are incrementally upgraded to provide better data transmission capabilities. The next section outlines the current state of data communications within GSM, and how they are planned to be extended to eventually become UMTS. The road ahead for other cellular radio network architectures will be very similar to that of GSM.

9.1 The Road to 3G Data

Current data services are constrained by the 9,6 Kbps limit on transmission speed on GSM radio links. This transmission rate is however already being used to provide data services to network users.

PC cards are used as interfaces between PC's and mobile phones. They connect PC's to data networks via mobile phones. The GSM network is digital. Data therefore doesn't have to be modulated and demodulated as with data transmission on the public switched telephone network (PSTN). Data is transmitted from the PC to the handset through the PC card. The handset transmits the data received directly onto the GSM network [HC98]. Data transmission and fax applications are already in use [HC98].

The GSM standard includes a service called SMS (Short Message Service). This service provides a two way paging capability based on store and forward principles. A mobile station can send a text message of up to 160 characters to another mobile station. If the other station is not powered on, the message is stored on a SMSC (Short Message Service Centre) and then transmitted later [HC98]. This data service is used to great effect to provide notification services such as email notification, emergency messages, news updates, stock market updates, advertising, and any other service that makes use of short text messages [GSM].

9.2 High-Speed Circuit-Switched Data Service - HSCSD

The current limitation of 9.6 Kbps over the radio link is due to the fact that only one of the eight possible TDMA channels is used for data. This limitation has been alleviated slightly by using less error checking, which raises the transmission rate to 14.4 Kbps in parts of the network which can afford to use less error checking [HC98].

The obvious way to increase transmission rates is to use more than one channel for data transmission. Now in trials is a service called high-speed circuit-switched data service (HSCSD) [Rys99a]. This service combines two to four of the channels for data transmission to attain rates of 28.8 and 56 Kbps [Rys99a]. With data rates like these a whole new world opens to cellular radio networks. User will be able to connect to the Internet and their own dial up network servers through the GSM network. Because of the fact that the GSM network provides encryption as a standard service [Bar98], the implementation of Virtual Private Networks on GSM networks is a natural extension of data services already provided.

The provision of acceptable data transmission rates enlarges the list of services provided to users by network operators, to include data applications and networking options.

9.3 General Packet Radio Service - GPRS

The increased transmission speeds attained by HSCSD does however not solve the data transmission problems of cellular networks completely. The GSM network is circuit switched (SS7), and therefore a complete virtual circuit is set up every time data or voice is transferred from one point to another across the network.

Most of today's network applications are bursty by nature, and rely on sending small amounts of information at irregular intervals e.g. email. It is very inefficient to create a virtual network circuit every time a user transmits a small amount of data at a time. To address this problem a service called general packet radio service (GPRS) is in the process of being developed [Rys99b].

The idea is to allow users to connect to packet switched data networks (IP and X.25) via a separate connection than the voice network. The two services will run in parallel. A user will be able to log onto the GPRS network, and gain access to any IP or X.25 network [Rys99b]. When connected to the network, the user will only pay for the amount of data transferred, as opposed to the time spent connected to the network. The user will be able to accept incoming calls whilst busy with a data transmission [Rys99b].

A GPRS network will be able to use all eight of the TDMA channels, and will therefore be able to sustain data rates of up to 164Kbps if the external network it interfaces with can sustain these rates. In order to implement GPRS, an extra node needs to be added to the GSM network architecture. This node, called the gateway GPRS support node (GGSN), will form the interface between the Mobile station and the packet switching network (IP or X.25) [Rys99a].

9.4 Enhanced Data Rates for GSM Evolution - EDGE

EDGE is the natural extension of GPRS. It uses the same network layer as GPRS, but is based on a new physical layer [Rys99a]. The physical layer is designed to transfer data at extremely high rates. This service is considered to be a third generation cellular service. It is expected to deliver data transmission rates of up to 500 Kbps under very good conditions.

9.5 Wideband CDMA - WCDMA

This will be the final third generation cellular architecture. It will be based on CDMA technology, but will use the same data networking system as EDGE/GPRS.

10 Value Added Services

Cellular network operators have to be at the forefront of technology in order to survive in an extremely competitive environment. Network infrastructures must continually be upgraded, and new applications and protocols developed that provide novel services to customers. Industry analyst Andrew Seybold stated that not making maximum use of advances in data transmission technologies, "could be a significant windfall for network providers who are interested in obtaining and keeping new customers." (Seybold's Outlook, August 1998) [Nac99].

To this effect, network operators are developing value added services (VAS). These are data services which are provided to users over and above the standard network services of voice transmission. Three technologies are being developed in order to make the development of these services possible for network operators. These technologies are: wireless application protocol (WAP), SIM application toolkit (STK), and Bluetooth.

10.1 WAP

WAP is a protocol which is used to enable wireless devices to transfer information in a reliable and effective manner. The constraints of the wireless medium forced developers to create a protocol which allows for the low bandwidth and unreliability of the wireless medium to be used effectively. WAP has an optimized protocol stack, and uses markup language at the application layer - WML and WMLScript [WAP]. It uses UDP instead of TCP, due to the unreliability of the radio link. It has transaction and security layers which were included to enable the development of applications allowing secure business transactions [WAP].

The WAP model interposes a WAP gateway between traditional web servers and WAP clients. The gateway translates HTML into WML, and compresses it into a binary form which is used to save radio-link bandwidth. The client has a WAP browser which interpretes the compressed binary files.

10.2 STK

STK is is defined in GSM 11.14, and defines an interface through which the subscriber identity module (SIM card) and cellular handset can communicate. This enables applications to be loaded onto the SIM card, and then run via the interface provided by the handset. The reason for wanting this type of functionality is that network operators need a computational platform to run data applications on telephone handsets. The obvious choice would be to use the handset itself. The problem is that all handsets don't provide the same computational platforms. Network operators however have complete control over the SIM cards they supply their customers with. It is therefore the ideal vehicle to implement and distribute their own applications through. The SIM card also provides built in security which is useful for business applications.

Applications will be available for download over the air from network operators, and users will be able to add and remove applications from the SIM card as needed. Sun has developed a Java card which runs Java programs and uses the STK interface.

10.3 Bluetooth

Bluetooth is a technology which allows mobile communication devices like handsets and mobile computers to communicate with one another without the use of cumbersome cables and software drivers [Nac99]. It is a joint effort by Ericsson, IBM, Intel, Nokia and Toshiba to make mobile data applications less cumbersome to use. A laptop PC will be able to communicate with a mobile phone, even if the phone is in another room. Used in combination with GPRS, a laptop will be permanently connected to a packet switching network through the mobile phone.

A user will only pay for data transmitted over the network. A mobile phone will be able to notify a laptop when an incoming email has been received. The user can then decide whether to pay for downloading the email to the laptop or not.

11 Conclusion

Cellular radio network infrastractures are growing at a tremendous rate. People are buying mobile phone contracts because they provide a useful and relatively cheap service. The services provided by network operators are being extended to include data services. Geographichal coverage of networks is enlarged by means of global roaming.

Cellular networks are in the process of becoming high speed data networks. The fact that the cellular user base is mobile, and that the user base is extremely large, makes the development of mobile data applications very lucrative. The development of these applications, have to take into consideration the nature of the cellular network, and will depend on the effective development of new communication protocols.

Industry is working on speeding up the development of mobile communication technologies. The ideal communicaton system where both voice and data services can be delivered regardless of location, network or terminal, is maybe not as far off as many think it is.

References

[AMH+99] I.F Akyildiz, J. McNair, J.S.M. Ho, J. Usunalioglu, and W. Wang. Mobility management in next-generation wireless systems. *Proceedings of the IEEE*, August 1999.

[Bar98] D. Barth. Privacy issues in mobile data communication. Internet - http://www.gsmdata.com/today_papers.htm, May 1998. Last Accessed: 28/02/2000.

[GSM] Gsm data knowledge site. Internet - http://www.mobiledata.com. Last accessed: 28/02/2000.

[HC98] K. Holley and T. Costello. The evolution of gsm towards umts. Internet - http://www.gsmdata.com/today_papers.htm, February 1998. Last Accessed: 28/02/2000.

[Nac99] S. Nachtsheim. Buetooth technology - the convergence of computing and communications. Internet - http://www.gsmdata.com/today_papers.htm, February 1999. Last Accessed: 28/02/2000.

[Rah93] M. Rahnema. Overview of the gsm system and protocol architecture. *IEEE Communications*, April 1993.

[Rap96] T.S. Rappaport. *Wireless Communications*. Prentice Hall, 1996.

[Rys99a] P. Rysavy. The evolution of cellular data: On the road to 3g. Internet - http://www.gsmdata.com/today_papers.htm, 1999. Last Accessed: 28/02/2000.

[Rys99b] P. Rysavy. General packet radio service - gprs. Internet - http://www.gsmdata.com/today_papers.htm, 1999. Last Accessed: 28/02/2000.

[Sco97] J. Scourias. Overview of the global system for mobile communications. Internet - http://www.gsmdata.com/overview.htm, October 1997. Last accessed: 28/02/2000.

[Tan96] A.S. Tanenbaum. *Computer Networks*. Prentice Hall, 1996.

[WAP] Wap forum. Internet - http://www.wapforum.com. Last Accessed: 28/02/2000.

[ZAB99] M. Zeng, A. Annamalai, and Vijay K. Bhargava. Recent advances in cellular wireless communications. *IEEE Communications*, September 1999.