1 Introduction

The human species is unique amongst all life forms in developing a sophisticated and rich means of communication – speech. While communication may have had its origins in the need for individuals to work co-operatively to survive, it is now deeply embedded in the human psyche and is motivated as much by social as business needs. Historically, this was met simply as individuals with similar interests and values chose to form small settlements or villages and all communication was face to face. It was not until the introduction of the telephone in the late nineteenth century that social and business networks could be sustained even when the individuals concerned did not live in the vicinity. Although the coverage and level of automation of the fixed telephony network improved dramatically over the next 100 years, the next major step, communication on the move, was only possible with the introduction of wireless networks.

The term ‘wireless network’ is very broad and, at various points in history, could have included everything from Marconi’s first transatlantic communication in 1901, the first truly mobile (tactical) networks, in the form of the Motorola walkie-talkie in use during the Second World War, to the wide-area private mobile networks in use by the emergency services and large companies since the late 1940s. However, ‘wireless networks’ didn’t really enter the public consciousness until the commercial deployment of cellular mobile radio in the 1980s. The planning and deployment of such networks to provide mobile voice and multimedia communications form the subject of this book.

1.1 Liberalisation of the communications industry

In the late 1960s and early 1970s the communications market was stagnant; the retail market *was* telephony, the only ‘applications’ being directory enquiries and the speaking clock! This was to change dramatically with the liberalisation of communications markets. The first significant instance of communications liberalisation occurred in the USA when the licence to compete for public-switched long-distance services was granted to MCI in 1969. This was followed in 1984 by the break-up of the AT&T monopoly into the long distance carrier AT&T and seven Regional Bell Operating Companies [1]. At the same time, in the UK, Mercury Communications was issued with a licence to build and operate a second network in competition with BT, the incumbent PTT. This

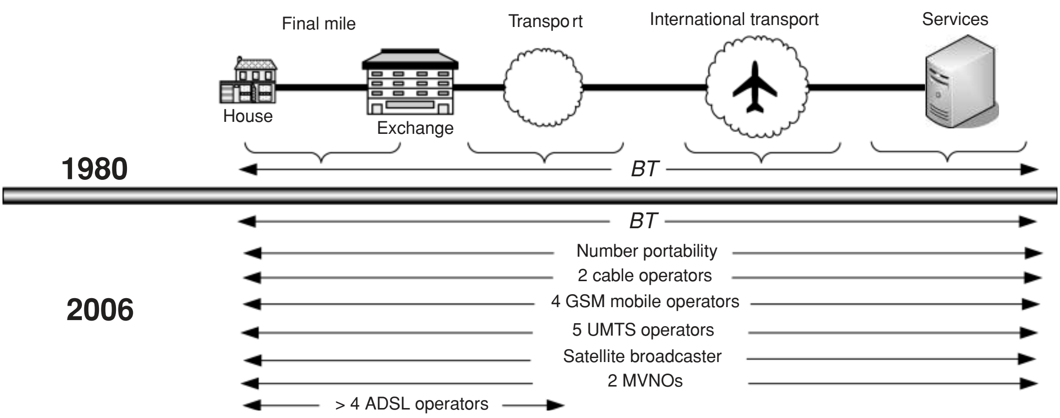

Figure 1.1 UK competitive environment – 1980 and 2006

process continued apace both in the USA and in the UK but it was not until the late 1980s that change started in the EU [2, 3], and culminated in the modification of an existing *Directive* to liberalise fully all European Union telecoms networks with effect from 1 January 1998. Since the 1998 *Directive* the EU has introduced a regulatory framework for the electronic communications sector [4], which is aimed at broadening the scope of competition policy. Figure 1.1 summarises other measures introduced in the UK during this period, which were generally adopted by mainland Europe in due course.

A key result of these changes was that current and future operators could, for the first time, provide almost any communication service; the previous silo structure of ‘fixed’, ‘mobile’, ‘cable’, etc., operators was eliminated. The impact of competition following liberalisation was not long in coming. For the first time, if consumers did not like the pricing or responsiveness of their current operators they could transfer to another service provider with little inconvenience. The result has been that prices for mobile voice have gone down to the point where they are little different from wired networks, and operators have had to look elsewhere to help sustain their margins.

1.2 Digitalisation of content

In parallel with the liberalisation activities of the last 30 years, technology has also moved forward dramatically. The inexorable increase of transistor density in integrated circuits, in line with Moore’s Law [5], has meant that the cost of processing and storing information in a digital format has, for most applications, become trivial. Computers, of course, have been conveying information, such as words and numbers, in a digital format since the 1950s but the first modern application conveying ‘content’ in a digital format was probably the use of the G711 codec [6] as part of the PCM-based telephone networks deployed in the late 1970s. This was followed by the compact disc in 1982, the DVD in 1996 and digital photographs with equal or better quality than their ‘film’ counterparts

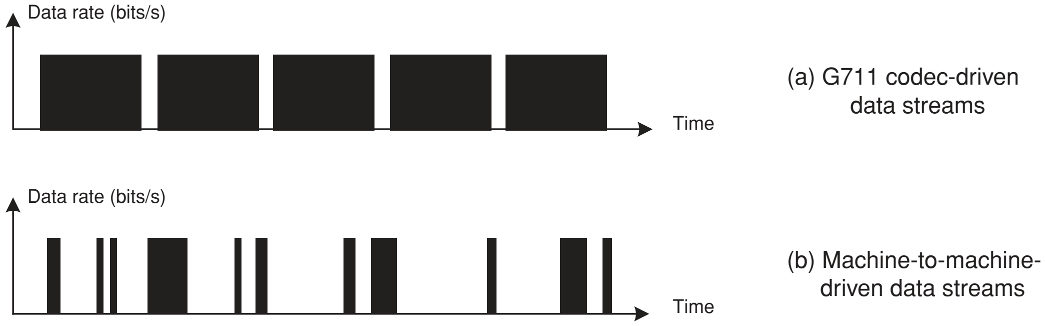

Figure 1.2 Application characteristics

in 2007. Paralleling the evolution of transmission has been the rise of new applications, such as multiplayer network gaming and web browsing. Nonetheless, for the first time, one (digital) transport solution can convey any content likely to be of interest to the user, provided the data rate, delay and bit error rate requirements can be met.

The dramatic reduction in the cost of signal processing has also enabled much more sophisticated codecs to be used, which reduce the amount of information to be transported whilst still allowing reproduction of the original content without significant degradation. The new algorithms usually reduce the information sent by a combination of *differential encoding* (only transmitting changes from the previous frame) and *tokenising* information (recognising, for instance, that the possible sequential time samples of speech are in fact constrained by the human larynx). This is very significant in that it changes the coded output from a regular sequence of constant data bursts of the type shown in Figure 1.2(a) to sequences of arbitrary size bursts of lower duty cycle shown in Figure 1.2(b). This ‘bursty’ traffic is also representative of ‘gaming’, web browsing and, indeed, most of the newer applications. Such applications are most efficiently conveyed on ‘packet networks’ as it is normally possible to convey the traffic from several users over the *same physical link*, which would previously have been dedicated to support one circuit-based application. This is achieved by interleaving other user traffic in the gaps between the bursts of the first user’s traffic.

It is also straightforward to evaluate the sort of sustained bandwidths required. Although it is unlikely that wide area mobile systems will be used to sell and download full length DVDs, selling individual CD tracks on the move to a multimedia phone is already becoming feasible. If a transaction time of 10 seconds is desired, and assuming that a 5 minute track is about 350 Mbits, then sustained bandwidths of 35 Mbits/s are needed.

1.3 Changes in spectrum management

Until the 1980s, there had been a consistent approach to spectrum management since spectrum was first used for communications purposes and its vulnerability to unintended interference was realised: particular frequencies would be designated for specific purposes and other users would not be allowed in the same band. This was a perfectly adequate approach until the rapid adoption of cellular phones, and other new wireless communications systems in general, meant that more radio spectrum was needed. To complicate matters, suitable spectrum for wide-area mobile communications is sensibly constrained to the region between 300 MHz and 3 GHz. At frequencies lower than this, the frequency bands that can be allocated to a system are too narrow to simultaneously support a useful number of users and reasonable data rates. At frequencies much higher than 3 GHz, radio propagation increasingly requires something close to line of sight between users, which is clearly not practical for 10–20 km cells. The 'last hurrah' for this conventional spectrum policy was the EU Directive of 1986 requiring the reservation of the 900 MHz band for GSM.

The first change to spectrum management was to retain the concept of dedicated bands for specific uses but to attempt to regulate demand using market forces. In the decade from 1994, the FCC in the USA [7] and regulators in Europe [8] raised a total of about $130 billion through a series of spectrum auctions. More recently, the FCC [9], Ofcom [10, 11] and the EU [12] have all been exploring a more flexible approach to spectrum management. The objectives for these new regimes, which are being progressively introduced, are essentially common to the three regulators:

- Ensure that the needs of all users are met,
- Maximise the economic benefits of spectrum,
- Promote the use of spectrum efficient technologies,
- Seek ways of making more spectrum available.

This new regulatory environment is much more relaxed and is characterised by:

- Market-driven allocation (spectrum auctions),
- Technology neutrality (subject to interference management),
- Only broad use designation (generally in line with the ITU),
- Freedom to trade some or all of the licensed spectrum with others,
- Ability to sub-lease spectrum for secondary applications.

The new regulatory environment has implications for future wireless systems. It will encourage operators to review at regular intervals whether it is commercially attractive to replace a deployed system with a new network that is more spectrally efficient. The ability to cram more bits/second into the same spectrum allocation translates to more revenue. Similarly, if it is possible to use the same *air interface* standard but in a variety of channel bandwidths with only minor changes to the equipment, it will offer operators economies of scale with corresponding cost benefits. These considerations account for the attractiveness of the *OFDMA* air interface and *MIMO* found in WiMAX and the planned UMTS LTE standard [13], which will be discussed in later sections.

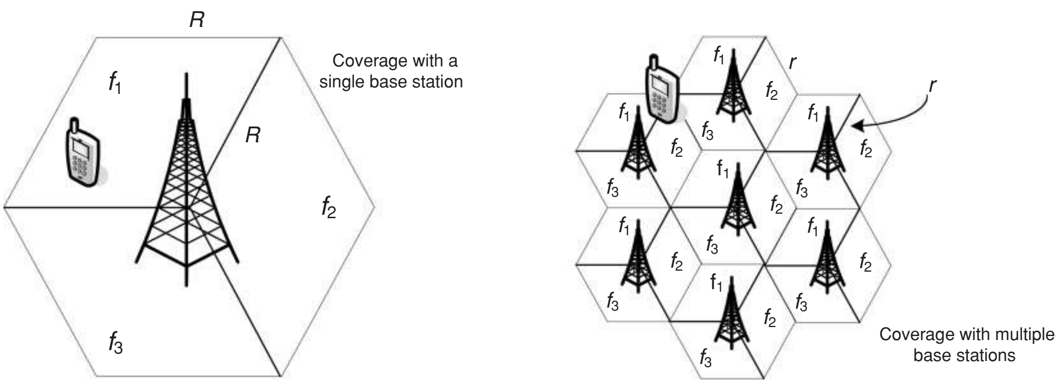

Figure 1.3 Why cellular?

1.4 Why cellular reuse?

The universal adoption of cellular solutions makes it apparent that such solutions are optimum for wide-area mobile networks. Why is this? Why not deploy one transmitter in the middle of the country, which is the configuration adopted for national time standards?

The primary consideration is coverage. Mobile radio systems need to support two-way communication, rather than unidirectional broadcasting. Whilst it is practical to achieve large coverage ranges in the downlink (from base station to the mobile) by using high transmit powers and large antennas, in the reverse direction (the uplink) the power is likely to be limited to 1 W or less if talk times of several hours are to be delivered by a small, light handheld device. A secondary consideration is that when networks mature, there is a need to address a much larger number of subscribers. Figure 1.3 shows that each deployment has three frequencies and is ‘sectorised’ into three cells. As the hexagonal shape factor and cell capacities will cancel out, the ratio of subscriber densities is simply the square of the ratio of the sides of the two hexagons, R/r. This will be discussed further in Chapter 3.

1.5 The drive towards broadband

Bringing together the discussion of these last few sections, it is apparent that the most efficient (and, therefore, most commercially valuable) operator would be one providing broadband packet-based access to content and applications over the transmission medium, which is most cost effective at the user’s current location. It is becoming clear that the current generation of ‘mobile’ operators will seek to sustain their profitability in the face of the eroding premium on mobile voice by offering a ‘one-stop shop’ for all the users’ telecommunication needs. The industry is already seeing mergers and acquisitions to enable operators to provide a common content and application base across TV, Internet, fixed and mobile systems – the so-called 'quadruple play'. This is expected to prove attractive to end users both from the perspective of a discounted offering for adopting the package (rather than individual services) and because of the integration of communications and content charges in one bill. For the operator, it provides access to new high-margin revenue and also reduces the likelihood of churn (with its associated cost) through customer inertia at the prospect of finding and negotiating new contracts with other carriers. This same inertia also provides opportunities to raise margins over a period of time. One can also anticipate a new and more flexible charging structure, based on a mixture of one or more metrics such as 'convenience', 'content value', 'quality of service' (QoS), 'bits shipped', etc.

Given that mobile broadband will be a major element of any operator's wide-area multimedia offering, what are the key elements of such a solution? For wide-area coverage, a cellular solution is essential for the reasons already discussed. However, what determines the applications that can be supported in the cell? A key consideration is the maximum data rate that can be sustained over the radio link to the user. This is governed by a fundamental relationship known as Shannon's Law [14]. Claude Shannon was a mathematician who helped build the foundations for the modern computer and developed a statement on information theory that expresses the maximum possible data speed that can be obtained in a single data channel. Shannon's Law makes it clear that the highest obtainable error-free data speed, expressed in bits per second, is a function of the bandwidth and the signal-to-noise ratio. It can be expressed as:

$$C = B \log_2(1 + S/N), \tag{1.1}$$

where:

C is the channel capacity in bits/s,
B is the channel bandwidth in Hz,
S is the signal power in the channel,
N is the noise power in the channel.

If this result is considered in the context of any system where mobile coverage is provided by a central base station, the power density from the central station must fall off as the distance r from the base station increases. Because the noise power N is independent of the distance r and constant, it follows that the maximum data rate available to a user must decrease as the user moves away from the base site. In a cellular system, where there is interference (arising from the frequency in the serving cell being reused in surrounding cells) it is necessary to substitute 'N' with '$N + I$' where I is the interference at the mobile location. If an application is to be supported over the whole cell, then it is clear that the system must be designed to support the required application data rate at the worst-case cell-edge location. Secondly, the total data rate (or throughput) provided by the cell is the summation of the data rate supported by each mobile at its location. Thus the cell capacity and the peak data rate at the cell edge are both ultimately limited by the spectrum available to the operator. It is primarily these two considerations and the finite amount of spectrum suitable for mobile purposes that account for:

- The large fees paid in auctions for mobile spectrum,
- The continued evolution of air interface designs in efforts to get closer to the theoretical Shannon capacity limit.

1.6 Organisation of the text

Chapter 2 seeks to establish a common understanding of the way most cellular networks operate, using the ubiquitous GSM system as a baseline, and discusses the key differences that can be expected in networks providing fixed or ‘nomadic’ wireless access. The concept of ‘contexts’, central to packet services in GSM and UMTS, will be covered in some detail. Chapter 2 will also explore the factors that contribute to cellular network operating expense, to determine activities that heavily impact the operator profit and loss account. Finally, the profit and loss account is used to establish metrics that can be used to assess the relative merits of wireless network technologies addressed in later chapters.

Chapter 3 develops the principles and processes that are used to plan wireless access networks. The major focus will be on cellular networks, as these usually represent the most complex planning cases, but the corresponding process for *Wi-Fi* will also be discussed. Circuit voice networks will be examined initially, but the treatment will be extended to understand the additional considerations that come into play as first circuit data and subsequently packet data based applications are introduced. Significant space is dedicated to the way in which packet networks are planned to deliver guaranteed QoS and thus ensure satisfactory operation of arbitrary applications. With the planning sequence understood, the way in which information from such processes can be used to explore the potential profitability of networks well in advance of deployment is addressed. Choices regarding which applications are to be supported in the network and the quality of service offered are shown to have a major impact on the profitability of projects.

Chapters 4 to 8 develop a *detailed* radio access network (RAN) design process to complement the generic principles and processes discussed in Chapter 3. These chapters will address, step by step, the practical planning process.

Chapter 4 describes the steps in the detailed planning process that are essentially common regardless of the specific air interface in use. It takes cell sizes and estimates of the numbers of cell sites from the generic planning activity described in Chapter 3 and, using *specific* topographical data for the regions to be planned, selects *actual* cell locations, antenna heights, etc., for the network.

Chapters 5 to 8 deal with the multiplicity of aspects that are specific to the planning of a particular radio access network (RAN). A major part of these sections will be dedicated to providing a detailed understanding of the particular factors that influence the coverage, capacity and system latency for each of these air interfaces. It is an understanding of such factors that enables system designers to exploit fully the potential of each air interface. The following air interfaces are addressed in the four chapters:

- GSM/GPRS/EDGE,
- UMTS,
- OFDM,
- 802.11 mesh networks.

Each air interface chapter concludes with a planning example.

Chapter 9 provides an overview of the system architecture, principal network elements, transmission and service support for two distinct core networks. The first architecture reflects the networks that would have dominated deployments up to 2006/7, when the majority of traffic was circuit-based voice and, with most services, supported off the MSC. The second architecture is becoming increasingly important and will come to dominate deployments in the future. It features IP transport for all services with applications hosted by IMS. The chapter concludes with the dimensioning of IP transmission required to transport multiple applications with different but guaranteed QoS.

Chapter 10 discusses the last phases of the network life-cycle: provisioning, operation and optimisation. The activities to be performed in the provisioning and operational phases are described in a technology-agnostic manner as far as possible, whereas methods and techniques for radio network performance optimisation are explained separately for the GSM, GPRS and UMTS access technologies. In addition to identifying the sources of information and discussing analysis techniques used for optimisation, indications of the expected values of the key performance indicators (KPIs) of optimised systems are provided.

References

1 V. Mayer-Schoenberger and M. Strasser, A closer look at telecom deregulation: the European advantage, *Harvard Journal of Law and Technology*, **12** 3 (1999).

2 T. Kiessling and Y. Blondeel, The EU regulatory framework in telecommunications – a critical analysis, *Telecommunications Policy*, **22** 7 (1998) 571–592.

3 UK DTI, *Communications Liberalisation in the UK, Communications and Information Industries Directorate* (UK Department of Trade and Industry, March 2001).

4 The European Commission, *Overview of the EU Regulatory Framework for the Electronic Communications Sector*, http://europa.eu.int/information_society/policy/ecomm/todays_framework/overview/index_en.htm

5 G. Moore, Cramming more components onto integrated circuits, *Electronics*, **38** 8 (1965).

6 The ITU, G 711 codec, www.itu.int/rec/T-REC-G.711/e.

7 P. Cramton and J. A. Schwartz, Collusive bidding: lessons from the FCC spectrum auctions, *Journal of Regulatory Economics*, **17** (2000) 229–252.

8 P. Klemperer, How (not) to run auctions: the European 3G telecom auctions. *European Economic Review*, **46** 4–5 (2002) 829–845.

9 FCC Office of Engineering and Technology, *FCC Issues Guiding Principles for Spectrum Management*, Report No ET 99–6 (1999).

10 UK Ofcom and Radio Communications Agency, *A Statement on Spectrum Trading: Implementation in 2004 and Beyond* (2004).

11 UK Ofcom presentation, *The Spectrum Framework Review* (2005).

12 Communication from the Commission to the Council, *A Market-based Approach to Spectrum Management in the European Union (COM(2005) 400 final)* (2005).

13 *3GPP, UTRA-UTRAN Long Term Evolution (LTE) and 3GPP System Architecture Evolution (SAE)*, www.3gpp.org/Highlights/LTE/LTE.htm.

14 C. E. Shannon, *The Mathematical Theory of Communication* (University of Illinois Press, 1949) (reprinted 1998).

2 Wireless network systems

The discussion in Chapter 1 indicated that dividing the planned coverage area into a number of radio cells results in a more spectrally efficient solution with smaller and lighter end-user devices. In such a network, as the individual moves further away from the cell to which he or she is currently connected, the signal strength at the mobile eventually falls to a level where correct operation cannot be guaranteed and the call may ‘drop’. However, because the cellular system is designed to ensure good coverage over the plan region there will be one or more other cells at this location that *can* be received at adequate signal strength, provided some mechanism is found to ‘hand over’ the call to one of these cells. Most of the complexity in practical cellular systems arises from the need to achieve this handover in a way that makes this process as imperceptible to the user as possible.

This chapter aims to establish a common understanding of the way most cellular networks operate, using the ubiquitous GSM system as a baseline, and highlight the key differences that can be expected in networks providing fixed or ‘nomadic’ wireless access. It will also explore the factors that significantly contribute to cellular network operating expense and thus determine activities that impact the operators’ profit and loss account. Finally, the profit and loss account will be used as an agenda to identify wireless network technologies that are likely to change in the future.

The standards that define cellular networks, as a relatively new technology, are subject to constant evolution as vendors and operators work to improve performance and introduce new features. Even though GSM was first deployed in 1992, it is still subject to very active development and so it is necessary to define a particular version as the baseline for this discussion. The specific version of GSM known as GSM *Release 98* will be used for all discussion in this chapter; it supports circuit and packet transport along with much of the functionality found in more modern systems.

2.1 Cellular networks

At the risk of stating the obvious, the original requirement for cellular networks (and still the dominant source of revenue) was to enable people to communicate by voice whilst they were away from their ‘normal’ fixed-line home or office phones. The architecture of first-generation digital cellular systems, such as GSM, was, therefore, based on the