Mobile Video Telephony

For 3G Wireless Networks

David J. Myers

McGraw-Hill
New York Chicago San Francisco Lisbon London Madrid
Mexico City Milan New Delhi San Juan Seoul
Singapore Sydney Toronto

The McGraw-Hill Companies

Library of Congress Cataloging-in-Publication Data

Myers, David J.
Mobile video telephony : for 3G wireless networks / David J. Myers.
p. cm.
Includes bibliographical references and index.
ISBN 0-07-144568-4
1. Video telephone. 2. Mobile communication systems. 3. Global system for mobile communications. I. Title.
TK6505.M94 2004
621.3845′6—dc22 2004055977

Copyright © 2005 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

1 2 3 4 5 6 7 8 9 0 DOC/DOC 0 1 0 9 8 7 6 5 4

ISBN 0-07-144568-4

The sponsoring editor for this book was Stephen S. Chapman and the production supervisor was Sherri Souffrance. It was set in Century Schoolbook by International Typesetting and Composition. The art director for the cover was Handel Low.

Printed and bound by RR Donnelley.

This book is printed on recycled, acid-free paper containing a minimum of 50% recycled, de-inked fiber.

McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please write to the Director of Special Sales, McGraw-Hill Professional, Two Penn Plaza, New York, NY 10121-2298. Or contact your local bookstore.

Information contained in this work has been obtained by The McGraw-Hill Companies, Inc. ("McGraw-Hill") from sources believed to be reliable. However, neither McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw-Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that McGraw-Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Chapter

1

Migrating to Third-Generation Mobile Networks

Commercial mobile telephony networks have existed for over 20 years. This chapter traces the evolution of mobile telephony since its introduction and the reasons for its development, from the earliest analog networks through to the latest third-generation (3G) systems. The chapter concludes by looking at the ways in which conversational video telephony can be deployed on 3G networks.

1.1 2G Mobile Networks

The first-generation mobile services that were deployed in the 1980s were based on analog techniques. They suffered from relatively poor voice quality and frequent dropouts. A variety of signaling and coding schemes were used, limiting the ability of customers of one mobile network operator to make calls to another, especially between countries. Handsets and network equipment were expensive. As a result they did not become mass-market consumer services but were limited to businesses and other customers who could justify the high cost.

The introduction of standardized second-generation (2G) digital technologies such as *global system for mobile communications* (GSM) brought costs down significantly, making mobile services more affordable, in part due to economies of scale. Technology improvements reduced handset sizes and improved battery life. The introduction of prepay services, combined with the ability to roam across networks and access supplementary services such as voice mailboxes, resulted in rates of growth of customer numbers across the world that were probably unprecedented for any other new product or service. Mobile telephony

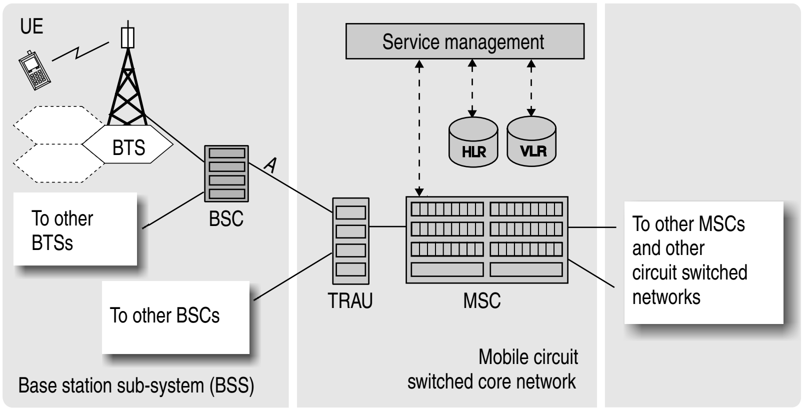

Figure 1.1 GSM 2G mobile network.

is claimed to have grown faster following the introduction of 2G systems than the Internet or television, services that also showed phenomenal rates of growth.

GSM was created by the European Telecommunications Standards Institute (ETSI). It is the dominant world standard for 2G networks, serving over 60% of all mobile customers worldwide. An indication of the impact of widespread standardization can be obtained by comparing the mobile market in Europe with that of the United States, where a range of incompatible 2G mobile technologies are used by different operators. Growth in the number of mobile telephony users in the United States has not been as strong as elsewhere and the ratio of mobile phones to people is still lower than in Europe.

A 2G network based on GSM is shown in Figure 1.1. This includes *base station subsystems* (BSSs) that are connected to the mobile core network. The mobile core network has interfaces allowing it to route calls to customers of other mobile operators and to customers of fixed network operators. Although these are not shown in Figure 1.1, it may also in principle have interfaces to packet networks, allowing its customers to make and receive calls to people using *voice over Internet protocol* (VoIP) on the Internet. In this case gateways would be required, network-based equipment that converts between the signaling protocols and speech coding formats used in the circuit-switched mobile network and those used for VoIP.

1.1.1 The GSM base station subsystem

Figure 1.1 shows that the BSS is composed of *base transceiver station* (BTS) units and *base station controller* (BSC) units. Each BTS covers one or more of the hexagonal cells that provide geographical coverage for mobile customers, and each BSC manages two or more BTS units. The BSC is connected to the mobile core network via a *transcoding and rate adaptation unit* (TRAU). Although the TRAU is not shown in Figure 1.1 as being part of the BSS, its function will also be discussed here.

The BTS provides the air interface allowing digital voice data to be sent to and received from mobile telephones, which are sometimes referred to as *user equipment* (UE). The mobile telephone contains a coder and decoder (codec) that produces a compressed digital speech signal according to one of the GSM Full Rate (GSM-FR), Half Rate (GSM-HR), or Enhanced Full Rate (GSM-EFR) speech coding standards, which result in digital speech signals at 13 kbit/s, 5.6 kbit/s, or 12.2 kbit/s, respectively. Each digital speech call is allocated a 16 kbit/s digital channel between the UE and the BTS.

The BSC manages two or more BTS units. It relays signaling and call data, deals with handover functions when mobile users move between cells, and raises alarms if problems occur. Because the mobile digital voice channels each require 16 kbit/s, four mobile channels can be multiplexed in a single 64 kbit/s channel or PCM timeslot between the BTS and BSC or between the BSC and the TRAU via the circuit orientated A interface shown in Figure 1.1.

The purpose of the TRAU is to convert between the 16 kbit/s digital mobile voice channels used in the BSS and the 64 kbit/s channels used in the mobile core network and in traditional fixed telephone networks. To do this it has to convert or transcode between the compressed GSM speech format and the G.711 *pulse code modulation* (PCM) speech format. If the TRAU is located at the opposite end to the BSC in the link between the BSS and the mobile core network, transmission costs associated with this link are reduced by a factor of 4 compared to locating it at the BSC end of the link.

1.1.2 The 2G mobile core network

The BSC is connected via the TRAU to a *mobile switching center* (MSC) in the mobile core network, which includes a number of MSC units interconnected with one another. Some of these MSCs, known as *gateway MSCs* (GMSCs), have interfaces to other switched networks for signaling, using *signaling system number 7* (SS7), and for transport of voice. The MSC is a traditional switch similar to a *central office switch* in a fixed telephone network.

The other components of the mobile core network shown in Figure 1.1 are the *home location register* (HLR) and *visitor location register* (VLR) databases. The HLR contains information of customer telephone numbers and the VLR that they are currently registered with. The VLR contains information of the customers currently served by the zone it is responsible for. Together with other components that have been consolidated for simplicity into the service management box shown in Figure 1.1, these are responsible for authenticating callers, determining the location of the customer they wish to call, setting up the call, and gathering data for billing purposes. The service management function also provides the *intelligent network* (IN) features that allow services such as rerouting callers to voice mail systems in the event that their desired destination is engaged or unavailable.

The switching infrastructure within the mobile core network for GSM-based 2G systems is based on circuit switching of PCM channels, using switching protocols such as SS7, in very much the same way as in fixed telephone networks. Dedicated 64 kbit/s bearers are set up between customers for the duration of the call. The bandwidth is further limited to 16 kbit/s within the BSS.

1.2 Evolution to 3G

After a sustained period of rapid growth, the mobile operators (and handset and equipment manufacturers) were faced with declining rates of growth as markets began to saturate. As a result they were keen to investigate possible new revenue sources. The success of the *short message service* (SMS), a 2G service allowing users to send and receive text messages up to 160 characters long that exploits existing signaling infrastructure, brought an unexpectedly significant source of revenues to mobile operators. It was a pointer to the future direction in which mobile telephony should evolve—the provision of data services as well as traditional voice telephony.

The data-carrying capacity of 2G mobile networks is limited. At best data can be transferred between the mobile device and the network at 16 kbit/s. In practice using wireless modems the maximum rate that can be achieved is 9.6 kbit/s. This is too low to provide an acceptable user experience. Standards have been produced for enhancements to existing 2G networks to provide some data capabilities based on existing infrastructure. These are usually referred to as 2.5G systems; they are expected in most, but not all, cases to be interim solutions in the run-up to full deployment of 3G.

The vision that drives 3G is that the mobile network will converge with the Internet allowing people to seamlessly access information and to communicate over different networks. High-speed digital services capable of being heavily used by a large number of customers need 3G mobile infrastructure to provide the bandwidth and data carrying capacity.

True 3G networks require additional spectrum allocation. Governments around the world auctioned licenses for this spectrum around 2000.

This raised a great deal of income for the governments concerned, at a high cost to the successful operators. This has delayed the launch of 3G services because the operators did not have sufficient money left over to invest in the required infrastructure. Operators now have more pressure on them to make sure that 3G provides compelling services that produce new revenue streams to recoup their investment in licenses, or to write off the cost of the licenses they bought.

1.2.1 Technology drivers

Since 2G networks were originally launched, a number of improved modulation and signal processing techniques have been developed. These increase the efficiency of spectrum usage at the air interface and provide a potential way of increasing capacity and overcoming the bandwidth and data-carrying limitations of 2G services.

Making use of this increased capacity to offer data services requires that the mobile core network be enhanced. This takes two forms, the upgrading of the existing circuit-switched infrastructure to extend 64 kbit/s channels all the way to the UE, and adding an overlay packet-based core network that operates in parallel with the circuit-switched mobile core to carry data traffic. This requires the functionality of the mobile network service management layer to be extended to support packet-based services. Customers using the packet interface need to be authenticated and billed as reliably as customers using the circuit-switched network. Wherever possible, existing Internet technologies and protocols are being used, and enhanced where necessary to take account of unique features of mobile networks. One example of this is that augmented location and authentication processes are required for data connections to take into account the mobility of customers.

The capabilities of mobile devices need to be extended to offer data services as well as voice services. Developments in this area have increased the computational power of handsets. This enables them to run the required network interfaces, protocol stacks, Internet browsers, and other application software without adversely affecting battery life or weight and therefore portability. Upgraded user interfaces have provided larger screen sizes with improved resolution and integrated cameras to allow video to be captured and displayed. Plug-in cards are becoming commercially available that provide mobile interfaces for laptops and other handheld computers such as *personal digital assistants* (PDAs).

These technology developments have been exploited by standardization groups, such as the *third-generation partnership project* (3GPP) to produce a roadmap and standards for 2.5G and 3G systems.

1.2.2 The evolution path to 3G

Mobile operators have invested large sums of money in their GSM infrastructure. They wish to get a return on this investment, and if possible reuse it where possible in providing 2.5G and 3G services. This desire has informed and impacted the development of the relevant standards, and has led to evolutionary rather than revolutionary approaches to migrating to 3G.

The initial stage in exploiting these technology improvements is the development of 2.5G capabilities to send and receive data, such as *general packet radio service* (GPRS). GSM networks can be upgraded to support GPRS using existing spectrum to provide a packet data interface with data rates that can peak at over 100 kbit/s. This involves upgrading the air interface of existing BTS units and providing BSC units with a packet-oriented interface to an overlay mobile core packet network. The service management layer of the mobile core network needs to be upgraded to provide support for the new services that are possible with GPRS such as the *multimedia messaging service* (MMS).

A further enhancement of GPRS is known as *enhanced data rates for GSM evolution* (EDGE). EDGE introduces a more advanced modulation scheme that allows data rates that are three times those of GPRS without requiring additional spectrum, and is seen as a viable alternative for operators who do not have a license for additional 3G spectrum allocation.

1.3 3G Mobile Networks

The 3G evolution path for GSM is to *universal mobile telephony service* (UMTS). UMTS 3G mobile networks deploy a new *radio access network* (RAN) in place of the BSS of GSM, as well as packet-based infrastructure in the mobile core network.

As we have already seen, the implementation of 3G networks is evolutionary. This process of evolution continues within the deployment of 3G networks, which is broken up into a number of phases. The first of these is usually referred to as Release 99 (after the year in which the recommendations were issued, 1999), though it is also sometimes known as Release 3. The Release 99 architecture includes a circuit-switched component, aimed mainly at voice services and a packet-switched component for data-based services. It is very similar to GSM augmented by GPRS/EDGE, except that the new RAN has greater capacity, provided by wider bandwidth and advanced modulation schemes.

The next phase in the deployment of 3G, known as Release 5, envisages that all traffic including voice will be carried over the packet network, using augmented *Internet protocols* (IPs).

In the short to medium term, however, many mobile operators will have a mixture of legacy GSM, 2.5G, and Release 99 3G networks. The

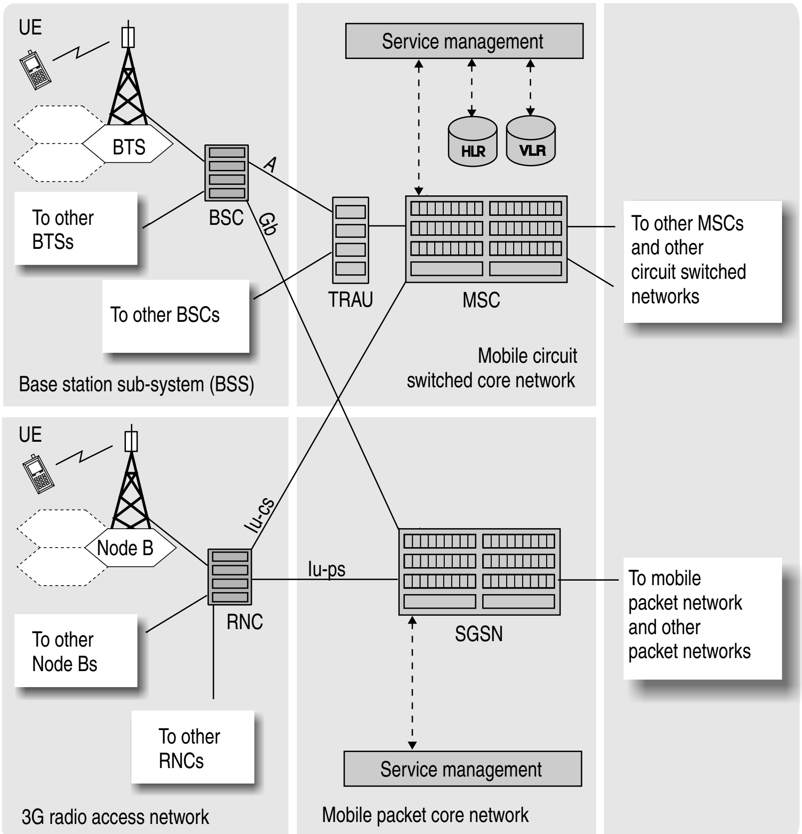

Figure 1.2 Evolved 2G GSM and 3G UMTS network.

network architecture for this setup is illustrated in Figure 1.2, which shows a 2G BSS and a 3G RAN, and a mobile core network split into circuit-switched and packet-switched subnetworks. The connection from the BSS to the packet core network via the Gb interface shown in Figure 1.2 allows 2G customers to access GPRS/EDGE-based services. The RAN connects to the circuit-switched core network via the Iu-cs interface, allowing the mobile operator to reuse 2G infrastructure to provide circuit-switched services to 3G customers.

1.3.1 The radio access network

The RAN shown in Figure 1.2 is made up of Node Bs and *radio network controllers* (RNCs). The Node B is equivalent to the BTS in a GSM BSS, and provides the air interface to the UE. The Node Bs are connected to RNCs, the equivalent of GSM BSCs. Unlike BSCs, the RNCs have direct connections to other RNCs that control Node Bs in adjacent geographical areas to provide soft handover as customers move from cell to cell. The RNC is connected to the circuit-switched mobile core network via the Iu-cs interface and to the packet mobile core network via the Iu-ps interface. The connections within the RAN and the interfaces to the mobile core network are all packet-based, using *asynchronous transfer mode* (ATM) technology, rather than using the circuit-based interfaces of a 2G BSS.

There is no equivalent to the GSM TRAU in the Release 99 3G network. For voice calls, speech is transported in its encoded form using the *GSM adaptive multirate* (GSM-AMR) standard operating in one of its eight encoding modes from 4.75 kbit/s to 12.2 kbit/s. Speech is not transcoded to G.711 and back when being transferred from one mobile customer to another. Speech destined for fixed circuit-switched networks will require conversion to G.711 at the interface between the mobile and fixed networks.

1.3.2 The 3G mobile core network

The MSC in a Release 99 network has equivalent functionality to the MSC in a 2G GSM network—the 2G MSC can be reused through a simple upgrade that includes replacing or augmenting the cards providing the circuit-based A interface with ATM cards to provide an Iu-cs interface, and to send and receive circuit-switched traffic at rates up to 64 kbit/s. Rates higher than this are possible, but for most proposed deployments of 3G networks a maximum rate of 64 kbit/s rate has been adopted—this matches with channels in the fixed circuit-switched network.

Data traffic to and from a UE is routed by the RNC over the Iu-ps interface to an IP switch/router known as a *serving GPRS support node* (SGSN), which is the entry point to the overlay packet-based core network. This performs authentication, allocation of dynamic IP addresses to the UE and IP switching and routing. Once a session is established incoming data traffic can be routed from the SGSN to the next point on the route to its destination via a *gateway GPRS support node* (GGSN), not shown in Figure 1.2. This may be to destinations in the fixed packet network, or to legacy-switched networks via a media gateway.

1.4 Mapping Video Telephony Services onto 3G Networks

In addition to continuing to offer speech telephony, 3G enables mobile operators to increase revenues by offering new services to their customers based on the delivery of digital information at high rates. Video telephony is one such service.

The two possibilities for offering video telephony on Release 99 3G networks are to use the circuit-switched mobile core network to deliver video telephony or to use the packet network.

Significant extra bandwidth is needed to transport video as well as speech. Release 99 3G networks effectively provide a 64 kbit/s circuit-switched path (even if it does make use of ATM between the MSC and the Node B). This is four times as much bandwidth as provided in the BSS of GSM networks, and means that if the speech is encoded at 12.2 kbit/s using the highest mode of GSM-AMR, there is around 50 kbit/s available for video. This is low compared to the bandwidth needed to transport uncompressed video, but video that provides an acceptable user experience can be obtained at this rate with advanced compression techniques. The reader should also bear in mind that the size of the screen that the video will be displayed on is of the order of 10 cm^2, and that head-and-shoulders scenes are likely to be the subject of the majority of mobile video telephone calls.

The alternative to circuit-switched mobile video telephony is to use the packet network. This may provide greater bandwidth, though bandwidth is not guaranteed in the way that it is when a dedicated circuit is available for the duration of the call.

3GPP has recommended packet-based video telephony for Release 5 and later 3G architectures based on the use of the *session initiation protocol* (SIP); however, for Release 99 the recommendation is for circuit-switched mobile video telephony, based on standards derived from the ITU-T H.324 recommendation, "terminal for low bit rate multimedia communication," and commonly referred to as 3G-324M.

The decision to choose circuit-switched 3G-324M is at least in part due to the relative immaturity of the packet core network components and protocols and the associated service management capabilities required. Customers are accustomed to a level of robustness and reliability for speech telephony services that they do not yet demand from the Internet. They will expect video telephony services to be equally robust and reliable. Mobile operators feel comfortable with the well proven and mature systems associated with speech telephony for capturing billing information, circuit-switched video telephony allows them to reuse these same systems. Until later releases of 3G networks when the packet network infrastructure is proven and mature, it seems to make good sense to take the circuit-switched approach.

In places such as Europe, Japan, Korea, Taiwan, and Australia 3G networks are commercially available and operators are offering video telephony services based on the circuit-switched approach. Handsets that support conversational video telephony on 3G networks are being produced by a number of major manufacturers. Japanese operator NTT DoCoMo has announced that its *freedom of mobile multimedia access*

(FOMA) 3G video telephony service, which is based on 3G-324M, passed the milestone of two million customers in January 2004 and three million customers just 2 months later. Perhaps the age of mass-market video telephony has arrived.

1.5 Chapter Summary

In this chapter we have reviewed the development of mobile networks since their introduction. The adoption of standards, advances in technology, the availability of extra spectrum, and the desire of operators to sustain their historical growth rates have driven this development. The latest 3G networks support higher bandwidths for both circuit-switched and packet-switched services, enabling mobile operators to offer a richer set of services.

We have seen that one of these services is mobile video telephony. This can be offered over either the circuit-switched or the packet-switched network. In the future it is likely to make use of the packet-switched network, but for reasons of maturity of technology circuit-switched video telephony services have been launched initially and are likely to continue, at least in the short to medium term.