

# The challenges of voice-over-IP-over-wireless

Göran AP Eriksson, Birgitta Olin, Krister Svanbro and Dalibor Turina

Much as second-generation radio access brought mobile telephony capabilities to the mass market, third-generation radio-access technology will introduce value that extends beyond basic telephony. The widespread growth of the Internet has created a mass market for multimedia and information services. The challenge of providing this access is twofold: from the market perspective, the challenge is to merge the installed base of users in cellular and Internet environments; in terms of technology, the challenge is to find common denominators for cellular solutions on one side, and efficient Internet access on the other. To successfully meet these challenges, the third-generation wireless systems must be designed to offer a multitude of services, provide considerable flexibility, structured QoS handling, and cost-effective access, while ensuring coverage with high radio spectrum efficiency.

These design challenges need to be resolved in the future evolution of UMTS and EDGE standards, beginning with the standardization of release 2000 in ETSI/3GPP, both regarding the radio-access networks and the common core network. A common denominator for this evolution is the Internet protocol suite, which provides end-to-end solutions for transport and network communication, including cellular (radio) access links over UMTS/EDGE.

The authors describe the requirements and current trends in the evolution of the third-generation wireless systems toward supporting more generic multimedia applications on a common platform. More especially, they discuss the challenges and solutions involved in designing a packet-switched radio-access network that efficiently supports VoIP service, highlighting session control for VoIP service, quality-of-cellular service for the cellular link, and the challenges of VoIP over wireless.

## Introduction

Today, the accumulated volume of data traffic is on the verge of surpassing the accumulated volume of voice traffic in all public networks. Given the growth in the areas of wireless voice and data, we see that the combination of mobile and Internet communication constitutes the driving force behind third-generation wireless systems, which promise to support at least 144 kbit/s (384 kbit/s) in all radio environments, and up to 2 Mbit/s in low-mobility and indoor environments.

The standardization of third-generation wireless systems is rapidly progressing in all major regions of the world. These systems—which go under the names of IMT-2000 (ITU), UMTS, and EDGE (ETSI/3GPP)—will extend the services provided by current second-generation systems (GSM, PDC, IS-136, and IS-95) with high data-rate capabilities. The main application for these services will be wireless packet transfer; for instance, for wireless access to the Internet. However, support will also be provided for

high data-rate circuit-switched services, such as real-time video.<sup>1</sup>

## UMTS

The universal mobile telecommunications system (UMTS) is being standardized in the Third Generation Partnership Project (3GPP), which is a joint effort between the European Telecommunications Standards Institute (ETSI) and the Association of Radio Industries and Broadcasting (ARIB, Japan). The basic radio-access technology for UMTS/IMT-2000 in all major areas of the world is wideband code-division multiple access (WCDMA). The 1999 release of the UMTS standard was the first to be implemented in commercial products.

The radio-access part—the universal terrestrial radio access (UTRA)—includes a frequency-division duplex (FDD) mode and a time-division duplex (TDD) mode. The FDD mode is based on pure WCDMA, whereas the TDD mode includes an additional time-division multiple access (TDMA) component.

The WCDMA system, which uses wideband direct-sequence technology (DS-SS), fully supports the UMTS and IMT-2000 requirements for 384 kbit/s wide-area coverage and 2 Mbit/s local coverage. Particularly noteworthy features of WCDMA are

- support for interfrequency handover, which is necessary for high-capacity hierarchical cell structures (HCS);
- support for capacity-improving technologies, such as adaptive antennas and multi-user detection.
- built-in service flexibility, which provides spectrum-efficient access for current as well as future applications; and
- efficient handling of bursty applications via an advanced packet-access mode.

WCDMA also provides efficient support for multimedia services; that is, for transferring multiple services on one connection.<sup>1,2</sup>

## EDGE

The GSM and TDMA/136 technologies make up the foundation on which the common radio access for data services will be offered. The *enhanced data rates for GSM and TDMA/136 evolution* (EDGE) concept, which ETSI and the Universal Wireless Communications Consortium (UWCC) have adopted as the migration path from GSM and TDMA/136, fulfills the requirements for third-generation wireless systems according to IMT-2000. EDGE is capable

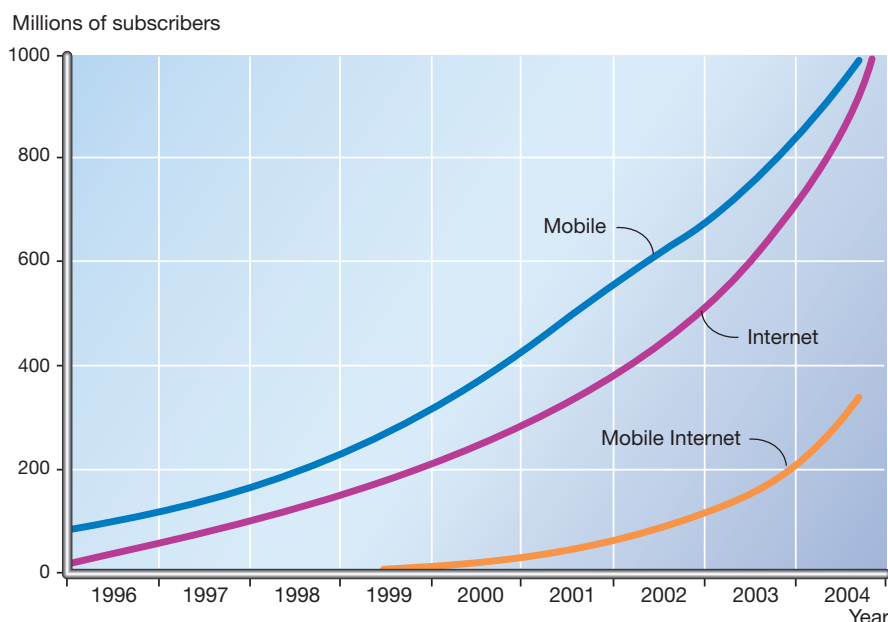
of offering data services of up to 384 kbit/s and is thus a global complement to the UMTS radio-access network.

The roadmap for EDGE standardization has been divided into two phases. Initial emphasis was placed on enhanced general packet radio service (EGPRS) and enhanced circuit-switched data (ECSD). According to the ETSI time plan, these standards were part of the 1999 release. The second phase of EDGE standardization, which is targeted for release in 2000, will define improvements for multimedia and real-time services. Other objectives will include the alignment of services and interfaces with UMTS, to allow EDGE and UMTS to share a common core network.

### Real-time IP applications over wireless

Second-generation radio-access technology brought mobile telephony to the market. Third-generation radio-access technology will extend beyond basic telephony: a common, IP-based transport and service platform will offer mobile users a multitude of real-time and interactive services.<sup>2</sup>

Typical services with real-time requirements are voice and video, as well as delay-sensitive applications, such as traffic-sig-

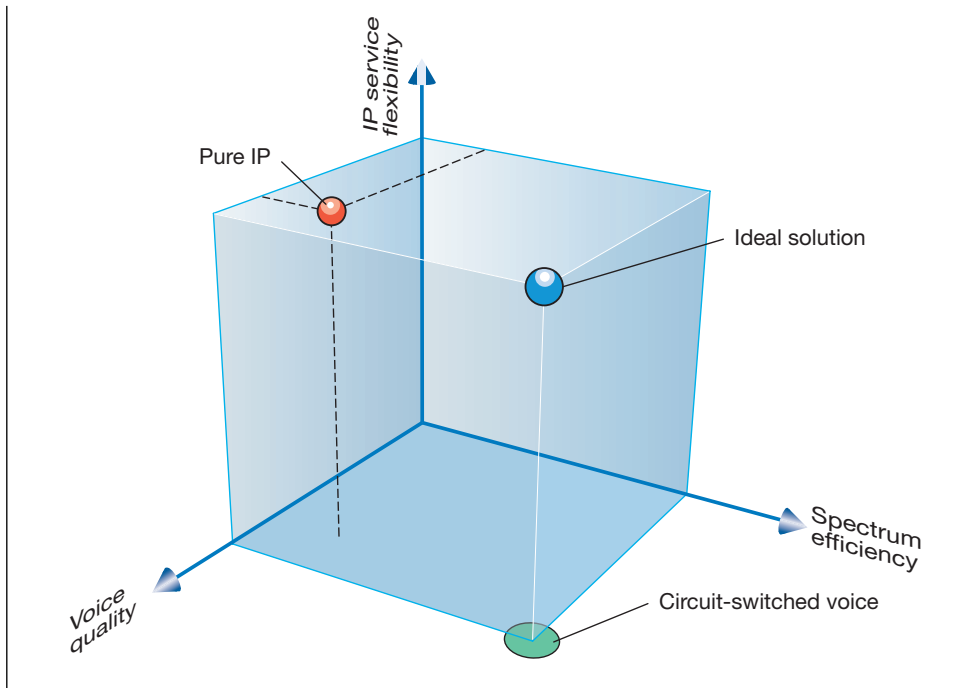


**Figure 1**  
The strong growth of mobile communication is expected to continue. Experts predict that by the year 2003/2004 there will be close to one billion subscribers of cellular systems worldwide. Similarly, the Internet will continue to grow. By 2004, the number of subscribers to the Internet is also expected to reach one billion. Of this group, more than 350 million persons will subscribe to the mobile Internet.

### BOX A, ABBREVIATIONS

|       |  |       |   |        |  |
|-------|--|-------|---|--------|--|
| 3GPP  | Third-generation Partnership Program               | H.323 | ITU-T standard for multimedia applications  | RRC    | Radio resource control                       |
| AMR   | Adaptive multirate                                 | HCS   | Hierarchical cell structure   | RSVP   | Resource reservation protocol                |
| API   | Application program interface                      | IETF  | Internet Engineering Task Force   | RTCP   | RTP control protocol                         |
| ARIB  | Association of Radio Industries and Broadcasting   | IMT   | International mobile telecommunication  | RTP    | Real-time transport protocol                 |
| BER   | Bit error rate                                     | IP    | Internet protocol   | SAP    | Service access point                         |
| CDMA  | Code-division multiple access                      | IPv4  | IP version 4  | SDP    | Session description protocol                 |
| CN    | Core network                                       | IPv6  | IP version 6  | SDU    | Service data unit                            |
| CRC   | Cyclic redundancy code                             | IS    | Intermediate standard   | SGSN   | Serving GPRS support node                    |
| CRTP  | Compression for real-time protocol                 | ISDN  | Integrated services digital network   | SIP    | Session initiation protocol                  |
| DS-SS | Direct-sequence CDMA                               | ITU-T | International Telecommunication Union – Telecommunications Standardization Sector | SMS    | Short message service                        |
| DTMF  | Dual-tone multiple frequency                       | LAN   | Local area network  | TCP    | Transmission control protocol                |
| DTX   | Discontinuous transmission                         | MS    | Mobile station (terminal)   | TDD    | Time-division duplex                         |
| ECSD  | Enhanced circuit-switched data                     | MT    | Mobile terminal   | TDMA   | Time-division multiple access                |
| EDGE  | Enhanced data rates for GSM and TDMA/136 evolution | NAS   | Non-access stratum  | TE     | Terminal equipment                           |
| EED   | Equal error detection                              | PDC   | Personal digital communication  | UDP    | User datagram protocol                       |
| EEP   | Equal error protection                             | PDCP  | Packet data convergence protocol  | UE     | User equipment                               |
| EGPRS | Enhances GPRS                                      | PDP   | Packet data protocol  | UED    | Unequal error detection                      |
| ETSI  | European Telecommunications Standards Institute    | PSTN  | Public switched telephone network   | UEP    | Unequal error protection                     |
| FDD   | Frequency-division duplex                          | PT    | Payload type  | UMTS   | Universal mobile telecommunications system   |
| FER   | Frame error rate                                   | QoS   | Quality of service  | UTRA   | UMTS radio access                            |
| GGSN  | Gateway GPRS support node                          | RAB   | Radio-access bearer   | UTRAN  | UMTS terrestrial radio-access network        |
| GSM   | Global system for mobile communication             | RAN   | Radio-access network  | UWCC   | Universal Wireless Communications Consortium |
| GW    | Gateway  | RLC   | Radio link control  | VoIP   | Voice over IP                                |
| GPRS  | General packet radio service                       | RNS   | Radio network server  | VoIPoW | Voice over IP over wireless                  |
|       |  | ROCCO | Robust checksum-based header compression  | WCDMA  | Wideband CDMA                                |
|       |  |       |   | WWW    | World Wide Web                               |

Figure 2  
The voice-over-IP-over-wireless challenge cube.



naling systems, remote sensing, and systems that provide interactive access to WWW servers. The focus of this article, however, is on voice service. The voice service of third-generation wireless systems must, at the very least, offer the same high level of voice quality, and be as spectrum-efficient, as present-day second-generation realizations. The challenge is to implement end-to-end service on IP-based transport.

The main advantage of running IP all the way over the air interface is service flexibility. To date, cellular-access networks have been optimized for voice quality and spectrum efficiency. The demand for service flexibility adds a new parameter, as illustrated by Figure 2. Since there are no dependencies between an application and the access network, almost anyone can develop new applications. But for services like voice over IP over wireless (VoIPoW), the main challenge is to achieve quality and spectrum efficiency.

To date, all cellular systems that provide voice service have been optimized in a two-dimensional space whose X-axis and Y-axis are voice quality and spectrum efficiency, respectively. Now, a third dimension is being added in the form of IP service flexibility. By bridging the radio interface with IP packets, we suffer a lot of protocol overhead,

which runs counter to the goal of spectrum efficiency.

## Network architecture overview

To facilitate our ensuing discussion, let us briefly describe the VoIP service. The basic components of the voice service are two user terminals with IP-based voice applications and a network that provides end-to-end transport between the terminals (Figure 3). The terminals exchange voice samples using the real-time transport protocol (RTP), which has been standardized by the IETF.

In some situations, terminals can establish and maintain communication without the involvement of a third-party entity. At other times, however, the two user endpoints cannot establish end-to-end communication without outside intervention; for example, when they do not know one another's IP address or do not use the same voice codec. In these cases, a control plane framework is used to route incoming traffic and to negotiate terminal capabilities (codec support, multiparty conferencing, and so on)—in traditional telecommunications this functionality, which is referred to as call control, is provided by, say, a GSM mobile switching center. In the IP world, there are

two main methods of providing call-control functionality: ITU-T Recommendation H.323, and the IETF session initiation protocol (SIP).

Originally intended for LAN environments, H.323 is an ITU standard for multimedia applications. Today, however, the standard is being adopted for broader usage. H.323 encompasses a complete architecture and a set of protocols, such as H.225 for call control and H.245 for bearer control. H.323 uses IETF protocols, such as the real-time protocol and the resource reservation protocol (RSVP).

Besides end-user terminals, the H.323 architecture encompasses gatekeepers, gateways, and multiparty units. In this context, emphasis is put on the gatekeepers and gateways—which constitute the VoIP server. The gatekeeper part is the controlling unit that provides call-control functionality; the gateway part contains the user plane functions. H.323 call control is based on Q.931, which is also used in GSM and ISDN.

The session initiation protocol, which is an IETF standard draft, is only one component in the IETF alternative to the H.323 paradigm for a complete multimedia architecture. Other necessary protocols and components include the session description protocol (SDP), the service access point (SAP), and the real-time control protocol (RTCP).

The session initiation and session description protocols (SIP/SDP) do not make up an architecture; they were designed for

session initiation. In contrast to H.323 and GSM/ISDN, SIP/SDP does not provide a complete call-control mechanism—an SIP proxy primarily provides routing and addressing services; device management is not included. However, the SIP proxy (or VoIP server) can be enhanced to include functionality for offering other services such as transcoding. The session initiation protocol is associated with a paradigm in which call control is distributed over several entities, and in which the user terminal plays a central role in coordinating these entities.

In summary, the two IP-based terminals exchange voice samples that have been encapsulated in RTP over the IP network. The terminals exchange control signaling between themselves or, with assistance from network entities such as a VoIP server, establish and maintain communication sessions through the network according to either the H.323 or SIP paradigm.

Both the session initiation protocol and H.323 support end-to-end solutions in which the network solely functions as a bearer. In this case we assume that an SIP or H.323 network call agent can, if so requested, support the end point (the terminal) with transcoding services.

The mobile terminal supports cellular access (UMTS/WCDMA or EDGE) and a complete VoIP application that is based on either SIP or H.323. We assume that an adaptive multirate (AMR) codec will be supported by future VoIP clients.

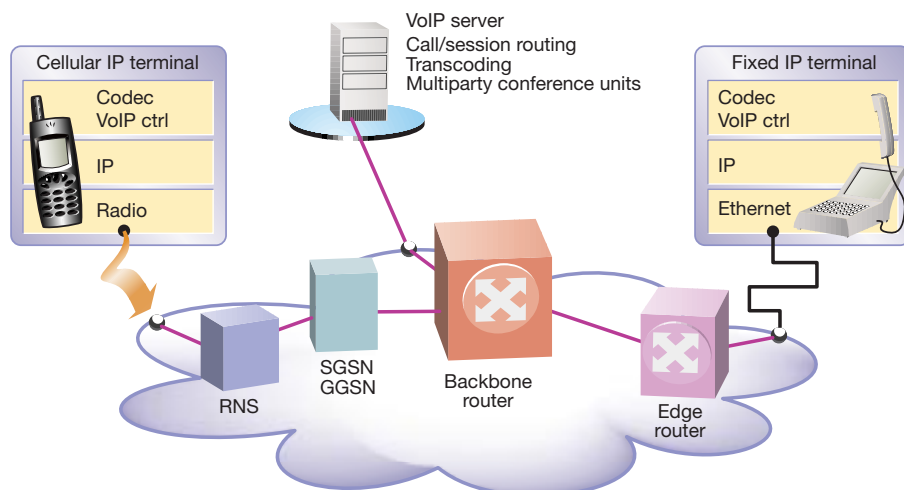


Figure 3  
Basic VoIP components.

In addition to basic UMTS packet-switched access, the network contains functions for adapting media, routing calls, and for authenticating users and services.<sup>3</sup>

## Overview of the QoS concept for the cellular link

The user's overall perception of quality of service (QoS) is a blanket judgement of the aggregate contributions from essential components of the communication system. In typical communication involving UMTS or GSM/EDGE, the transmission phase through the radio-access network (RAN) is only one part of the overall end-to-end communication. Thus, even when the radio-access network provides excellent quality of service, there is no guarantee that the end-user's perception of the service or application is good.<sup>4</sup>

### Bearer services within UMTS and GSM/EDGE

Given that frequency spectrum is a sparse resource in wireless communication systems, we can greatly benefit from applying a tailored-to-fit quality-of-service concept to the radio-access network. This way, using as few radio resources as possible, we can match each connection with the stipulated quality requirement.

To achieve a given network quality of service, a bearer service—with clearly defined characteristics and functionality—is set up from the source to destination. Each bearer service on a specific layer offers its individual characteristics through services that are provided by subordinate layers. The shaded sections in Figure 4 indicate bearer services that depend on services provided over the air interface.

Once again, because frequency spectrum is a sparse resource, we readily see the benefit of being able to classify traffic, in order to guarantee system capacity and quality of service. By being able to differentiate traffic flows in the network, we can define four application-related service classes within UMTS and GSM/EDGE:

- The *conversational service class* is used for real-time services, such as ordinary telephony voice; for instance, VoIP and videoconferencing. The vital characteristics of this class are low transmission delay and preserved time relationships, or low-delay variation, in the traffic flow.

- The *streaming service class* applies to real-time audio and video-streaming applications. In contrast to the conversational class, this category comprises one-way transport.
- Typical applications associated with the *interactive service class* are WWW browsing and Telnet. The fundamental characteristic of the interactive class is a request-response pattern, which makes round-trip delay an important characteristic. In addition, all data transfer must have a low error rate.
- The *background service class* is used for best-effort traffic. Examples of services in this class are electronic mail (e-mail), short message service (SMS) and file transfer. Here, too, all data transfer must have a low error rate, although the requirements for transfer delay are less stringent.

The transport for each service class can be configured in a way that optimizes the efficiency of the radio network and fulfills the quality-of-service requirement.

Different radio-access bearers (RAB) transport the services through the radio-access network. Each RAB is associated with a set of attributes that specify the required quality (bit rate, delay and error rate) and supply information on the characteristics of the traffic flow. This information is essential for

- providing a connection with good quality through the radio-access network; and
- using spectrum efficiently.

Examples of RAB attributes are service class, guaranteed bit rate, transfer delay, service data unit (SDU) loss rate, residual BER, and traffic handling priority.

### Service requirements for voice

#### *Unequal error detection*

Ordinarily, the bits in a frame from a cellular voice codec are divided into three classes: 1a, 1b and 2. Bit-error sensitivity varies between these classes—Class 1a includes the most sensitive bits; Class 2, the least sensitive bits.

In a typical second-generation system, the bits in Class 1a are covered by a cyclic redundancy code (CRC) that checks for errors in the frame. Thus we say that the voice frame uses an unequal error detection (UED) scheme.

If information on the different classes of bit-error sensitivity cannot be transferred from the codec to the radio-access network, or if the bits in the voice frame are not or-

ganized into classes, then the UED scheme cannot be used. Instead, an equal error detection (EED) scheme—a CRC that covers the entire voice frame—is introduced. To have the quality in these two cases, each must receive the same number of frames with a bad CRC.

Where circuit-switched traffic is concerned, only frames with a bad CRC contribute to the frame error rate. But in an IP-based radio network, frames with a bad CRC, frames that are lost due to jitter, and fatal errors in the IP header contribute to the frame error rate. By fatal errors in the IP header we mean user datagram protocol (UDP) checksum errors, errors in the link layer, and header decompression errors.

#### Unequal error protection

The calculation of the bit error rate (BER) only includes those errors that occur in bits not protected by the CRC. The residual errors in bits protected by the CRC must be as close to zero as possible. If any residual bit errors exist in the Class 1a bits, the voice decoder might produce noticeable artifacts.

If UEP is not available (but UED is), the lowest FER requirement of Class 1a and the BER requirement of Class 1b constitute the channel requirements.

#### End-to-end delay

The ITU-T recommends the following limits on one-way delay:

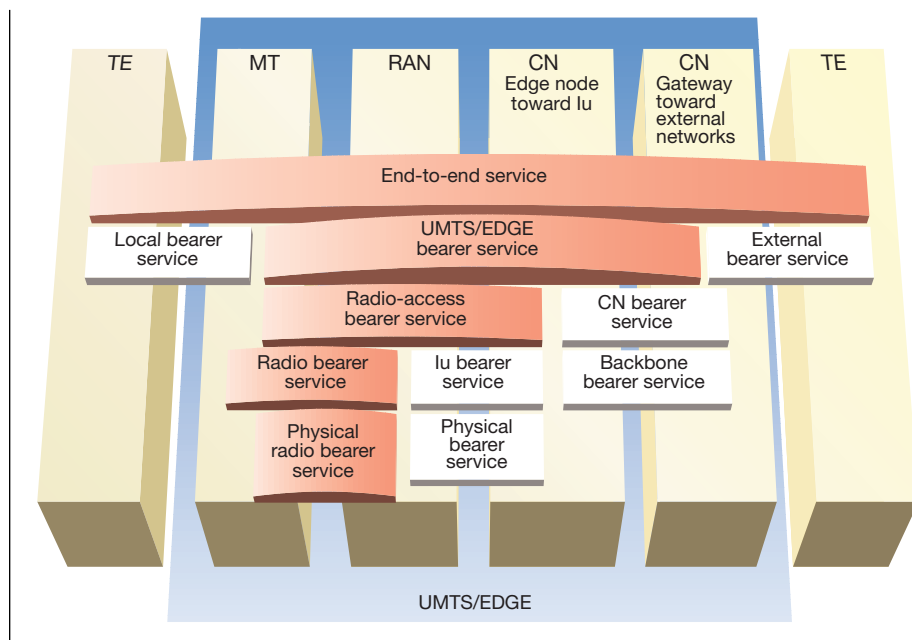
- 0-150 ms—acceptable for most applications.
- 150-400 ms—acceptable provided the administrator knows what impact the transmission time will have on the transmission quality of other user applications.
- Greater than 400 ms—unacceptable for general network planning; however, in some exceptional cases (such as satellite hops) this limit will be exceeded.

A typical second-generation cellular system has a one-way delay of less than 100 ms. Thus, to achieve the same high end-to-end quality, the one-way delay requirement for the conversational service (terminal-to-PSTN) of a third-generation wireless system must also be less than 100 ms.

## Optimizations for VoIPoW

### Challenges

The main objective of the VoIPoW concept is to port voice service to the new packet-data-based platform while retaining the



**Figure 4**  
Hierarchical structure of bearer services within UMTS and GSM/EGDE.

quality of service that users associate with present-day circuit-switched realizations (for instance, GSM circuit-switched voice service). To meet this objective, several quality-of-service issues have to be resolved on the core network and radio-access network sides. One objective is to offer voice as an integral part of emerging public and private conversational multimedia services, thereby converging the service with architecture solutions developed in the IP domain. A long-term migration solution to this realization implies that equivalent solutions to most telephony services must be offered over the wireless IP solutions.

It should be noted that the voice-only application is considered a special case—as part of a multimedia session, voice has more complex requirements; that is, in terms of BER and delay, different streams in a multimedia session may have vastly different QoS requirements. The Internet world on the fixed side is currently experiencing a tremendous growth in number of users and applications. At the same time, the goal of bridging the wireless air interface introduces additional design challenges. The main constraint is the capacity of the radio link, which is a sparse resource that must be used with care. Consequently, one additional requirement for the radio-access net-

work side is to obtain the same or comparable spectrum efficiency for voice services as from present-day circuit-switched systems.

To achieve spectrum efficiency, we can benefit from characterizing different packet data streams in terms of bandwidth and delay requirements. Characterizations of this kind are useful when implementing admission control algorithms that accommodate multiple user data streams in available spectrum. We can also benefit from applying different methods of minimizing the amount of data (such as RTP/UDP/IP header compression and session signaling compression) to obtain adequate spectrum efficiency for voice. By minimizing the amount of overhead, we obtain nearly the same level of spectrum efficiency as for the reference case, in which only voice frames are transferred across the radio through a circuit-switched connection.<sup>2</sup>

#### Traffic classification in IP networks

When a bearer is requested from the UTRA or GSM/EDGE radio-access network, the service it will bear is described by a number of parameters (radio-access bearer attributes), such as guaranteed bit rate, residual BER, packet loss rate, and delay. Therefore, to request an optimum VoIP bearer from the radio-access network, the requesting entity must know the attribute settings of the radio-access bearer. The procedure for determining and assigning values to the attributes of the radio-access bearer is called

*traffic classification*. Obtaining the information needed to set these attributes is a fundamental problem of transferring IP services over a radio network. In the IP world, the application and transport are separate and independent, whereas in the cellular world the two are usually integrated. Consequently, service data must be obtained for setting up an appropriate radio-access bearer. Apart from the RTP stream, the VoIP session also contains control signaling traffic flows with different characteristics. Moreover, a VoIP session might easily be extended to include video or a background FTP transfer.

#### Explicit method

The introduction of an application program interface (API) between the application and the radio link enables the explicit ordering of proper radio-access bearers. The API can be client-based or server-based. By client-based we mean that the API is between the application part and the radio part in the mobile terminal. By server-based, we mean that the API is defined on the fixed side of the network between the call agent (an H.323 server, for example) and the cellular radio network.

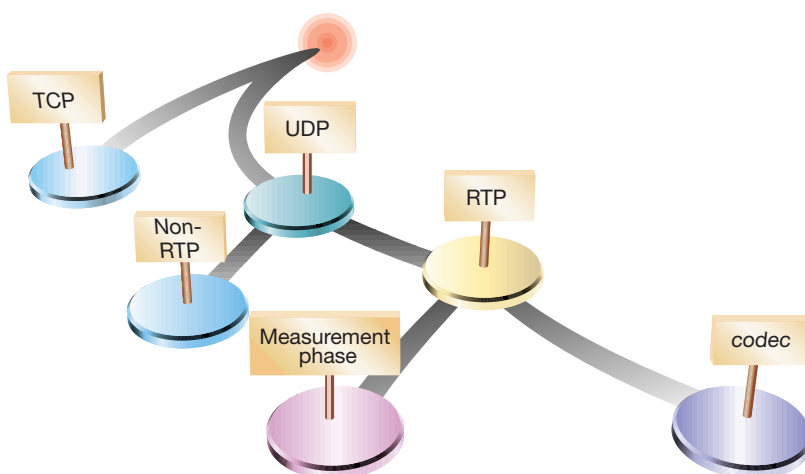
#### Implicit method

A more transparent method of obtaining service data is to develop a flow-classification algorithm that examines the headers of packets and extracts information on flows for identifying and characterizing them. Figure 5 shows the principle that applies in real-time situations. The RTP header contains a payload-type (PT) field that identifies the source codec. However, the PT can also be dynamic, in which case information on the codec is unavailable. In this case, the algorithm measures parameters, such as packet size and inter-arrival time. From these parameters, it might be possible to identify the codec or to give direct input to the radio-bearer parameters. How much information the algorithm can provide depends on the depth of the tree. For instance, the algorithm could be extended to capture specific signaling messages. Obviously, the complexity of the tree and the time it takes to go through it are limiting factors.

#### Header compression for real-time IP

The large headers of the protocols used when voice data is sent over the Internet constitute a major problem with voice over IP over wireless. An IP packet with voice data will have an IP header (20 octets), a UDP head-

Figure 5  
Flow classification tree.



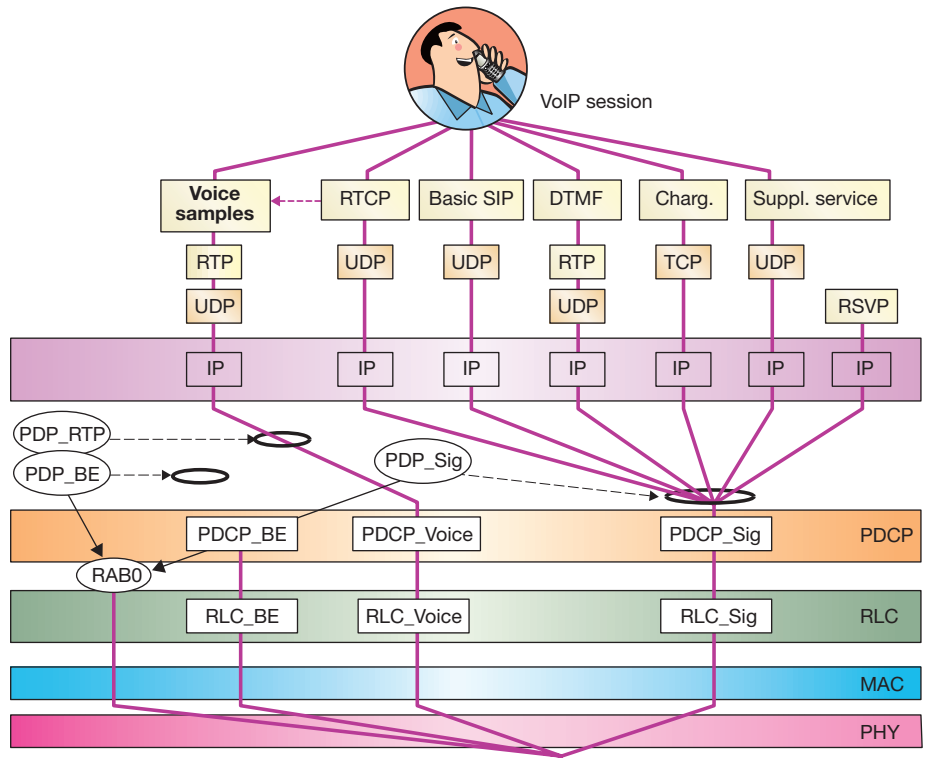


Figure 6  
Conceivable realization of a VoIPoW user session.

er (8 octets), and an RTP header (12 octets) for a total of 40 octets. In IPv6, the IP header is 40 octets, which raises the total to 60 octets. The size of the voice data depends on the codec and can be as small as 15 to 30 octets. These numbers make a strong argument for terminating the IP protocols before the air interface: the IP/UDP/RTP headers consume too much bandwidth and make inefficient use of valuable radio spectrum. However, it is possible to overcome this problem with header compression.

No header information in a voice packet can be discarded, but because there is a high degree of redundancy in the fields of headers of consecutive packets that belong to the same packet stream, this information can be compressed by means of header-compression algorithms. These algorithms maintain a context, which is essentially the uncompressed version of the last transmitted header at each end of the channel over which header compression occurs. The compressed headers solely carry changes to the context. Static header fields need not be transmitted

at all, and fields with only minor changes can be updated with only a few bits. However, when frames are lost or damaged over the channel, as can occur on cellular links, the context on the downstream side cannot be updated properly and the decompression of subsequent headers produces incorrect headers. Thus, the header-compression schemes must have mechanisms for installing context, for detecting when the context is out of date, and for repairing the downstream context when it is incorrect.

Several header-compression algorithms are being standardized by the IETF. For VoIP, the compressed RTP (CRTP) can compress 40-octet IPv4/UDP/RTP headers to a minimum of 2 octets.<sup>5</sup> To repair context, the CRTP relies on an upstream link over which the decompressor sends requests for updating the headers. All packets received by the decompressor are lost while the context is out of date, since the headers cannot be decompressed. The round-trip time over the link will thus limit the efficiency of the context-repair mechanism.

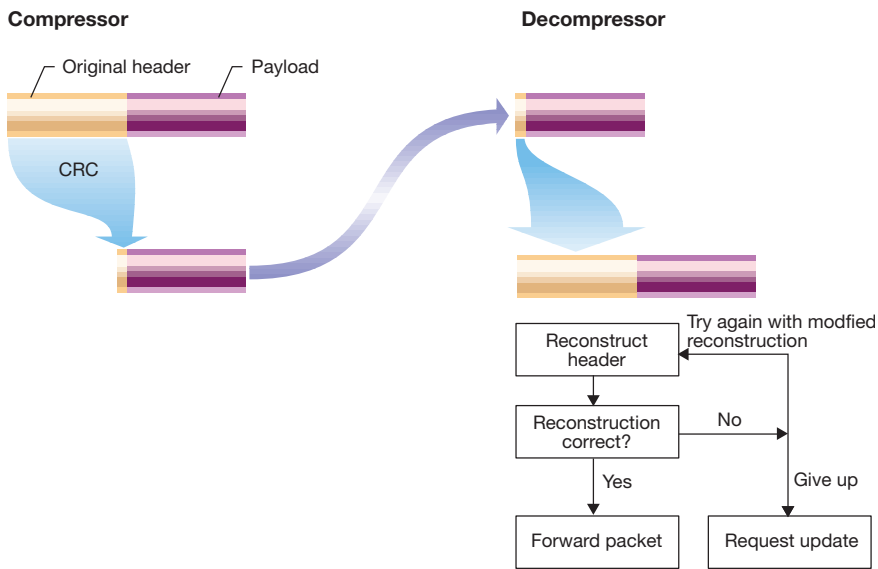


Figure 7  
ROCCO—the robust checksum-based header-compression scheme.

Link-level simulations show that the packet-loss rate for CRTP is about four times greater than for an ideal scheme in a VoIP-over-WCDMA scenario. To be viable, the header-compression scheme for VoIP-over-wireless cannot be less efficient but must be less fragile than CRTP.

Ericsson's response to these requirements is a header-compression scheme that provides a high degree of compression and is suitably robust for use in cellular applications (Figure 7). The robust checksum-based header-compression (ROCCO) scheme is heavily geared toward local context repair.<sup>6</sup> A checksum that covers the original (uncompressed) header is included in the compressed header to introduce a reliable way of detecting when

- the context is out of date; and
- local attempts to repair the context have succeeded.

Moreover, to achieve really good compression and robust performance without losing generality, Ericsson has also introduced a compression profiles concept. To achieve optimum performance, different compression profiles handle different RTP streams and channel conditions. Various profiles are being developed for voice and video streams. A general, all-purpose profile might also be developed. Figures 8 and 9 compare ro-

bustness and compression performance against channel quality for VoIPoW service when tested with different types of compression schemes and channel types.

In addition to the ROCCO header checksum, a code that is included in the compressed header provides the decompressor with information on how the header fields have changed; for example, due to loss over the cellular link. For the ROCCO VoIP profile, this code contains enough information on previous headers to enable local repair of the context after several (up to 26) consecutive packets have been lost between the compressor and the decompressor. The profile with the maximum compression ratio has a minimal header size of one octet. The ability to repair context locally virtually eliminates the negative effect of long round-trip times on header-compression performance.

#### Radio-access bearers for VoIP

The main challenge in designing radio-access bearers is to find suitable implementation points (Figure 2). These are dominated either by requirements for IP service flexibility or by requirements for spectrum efficiency. Different points satisfy these requirements to different degrees. For example, the amount of protocol overhead (that is, the IP header) cannot be minimized if IP service flexibility must be combined with end-to-end encryption. On the contrary, if we solely want to implement voice service, we can use a radio-access bearer to obtain a solution, which in terms of spectrum efficiency, is comparable to that of present-day solutions.

In summary, the radio-access network of a third-generation all-IP wireless system should support VoIP applications by means of a

- voice-optimized radio-access bearer whose service characteristics and spectrum efficiency match that of circuit-switched voice—for instance, by means of the AMR codec.<sup>4</sup> All IP-related overhead is terminated before bridging the wireless air interface. A trusted proxy is needed on the network side and no IP service flexibility is offered to the mobile user.
- voice-optimized radio-access bearer whose service characteristics and spectrum efficiency are similar to that of circuit-switched voice. Voice frames are transferred across the air interface together with the compressed IP overhead—for example, using header-compression algorithms.<sup>1</sup> Additional optimization might

be obtained from unequal error protection (UEP) or from unequal error detection (UED) that matches the voice codec in use. This solution allows for migration to an all-IP wireless system that efficiently supports telephony service.

- generic radio-access bearer for real-time user-to-user IP connectivity; that is, by supporting data streams that belong to the multimedia session and using RTP/UDP/IP. This solution provides total IP flexibility to new IP multimedia applications and in cases where security mechanisms (IPsec) do not allow for header compression.

### Overview of VoIPoW realizations in UTRA and GSM/EDGE networks

To provide efficient and cost-effective solutions for deploying end-to-end IP-based multimedia services that satisfy the elevated standard of QoS for UMTS (the radio part), the UTRA network must select the radio-access bearers that best suit the carried service. By optimizing how the radio resource is provided, we can also maximize system capacity. Figure 8 depicts a global overview of the data streams that relate to a VoIPoW call.

Data streams with similar characteristics and approximately the same QoS requirements can be logically grouped and mapped onto the same radio bearer. Within a VoIPoW session, a few relevant classes have been identified:

- RTP voice stream (RAB2)—the transfer of voice data requires the lowest possible delay, no jitter, and BER of  $10^{-4}$  (less if a robust header-compression mechanism cannot be used). Since low delay requirements do not permit retransmission, the preferred choice for the voice stream is a transparent RLC service—with guaranteed peak bit rate—that does not introduce any additional protocol overhead.
- Application signaling (RAB1)—RTCP, H.323 or SIP, RSVP. This category has less stringent delay requirements but demands better data integrity than voice. The requirement for data integrity can be met by means of retransmission at the RLC level. In some cases, the quality requested by this kind of signaling cannot be matched by a pure best-effort radio-access bearer. Therefore a guaranteed minimum bit rate should be considered within the acknowledged RLC service mode. To avoid over-providing orthogonal codes—provided more bandwidth is

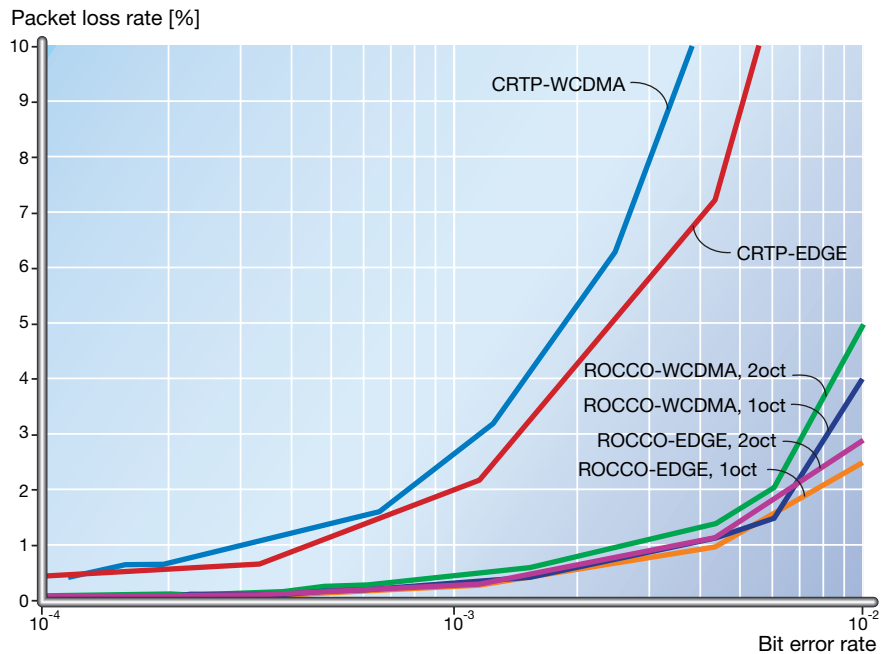
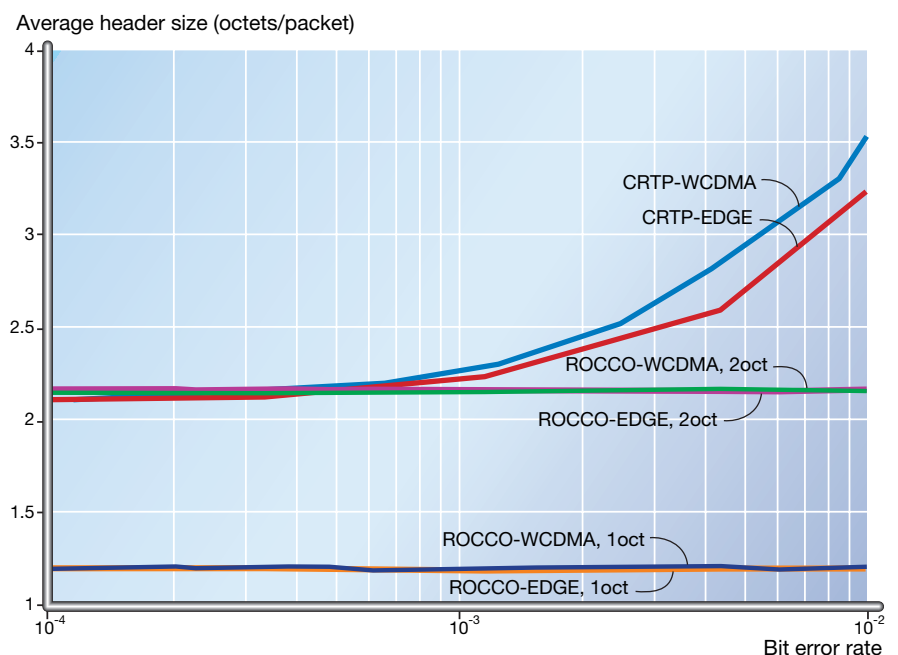
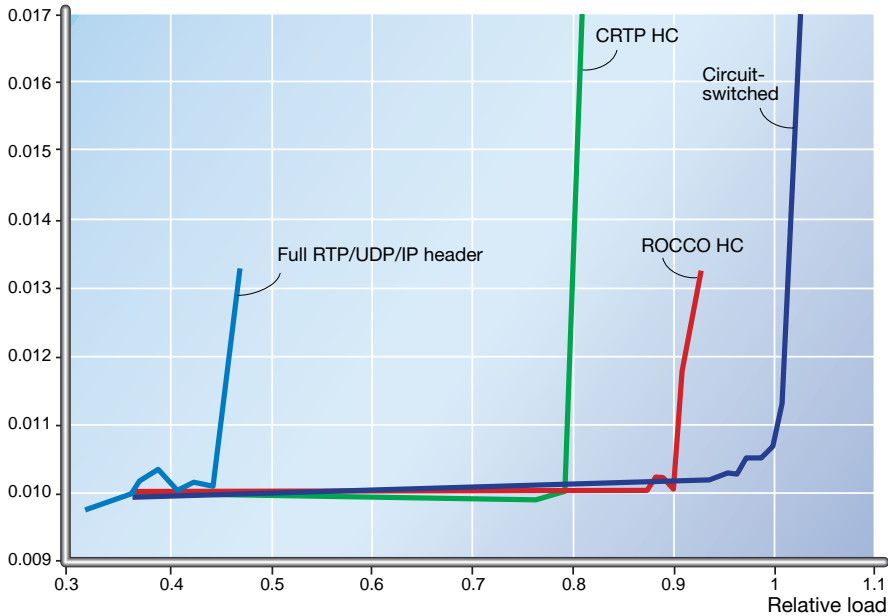


Figure 8 Global overview of data streams that relate to a VoIPoW call.

Figure 9 Simulated results of CRTP header compression.



System quality,  
the 95th percentile of FER



**Figure 10**  
Simulation results of UMTS voice services.

required for occasional bursty signaling during voice—the use of common or shared channels should be considered.

- Radio resource control (RRC) and non-access stratum (NAS) signaling (RAB0)—UTRA network control signaling is used to exchange messages between the UMTS network and the user equipment. Rapid and reliable delivery of messages can have a major impact on the performance of the entire system. Therefore, the transmission of UTRA network messages must be guaranteed, reliable, and have low delay and top priority.

Note: not all the data streams of the VoIPoW session depicted in Figure 8 are active at the same time. In fact, most of them can be decoupled.

### Evaluating the capacity of VoIPoW

System simulations have been made to determine how the introduction of IP service flexibility in the WCDMA-based UMTS system affects capacity. The focus of the simulations was on the radio link, which is considered to be the bottleneck of the system. Four different cases, each representing a different technical solution, were studied:

1. Circuit-switched voice service with equal error protection. This case is included as a benchmark. The capacity of this solution is used to normalize the other results. In all likelihood, the circuit-switched voice service in WCDMA will use unequal error protection. Thus, in terms of capacity, the studied results are slightly pessimistic. Nonetheless, the results of this solution suffice for comparison with voice-over-IP cases, to investigate the effects of the upper protocol layers on system capacity. For the air interface, this solution is the equivalent of one that terminates IP-related overhead in the network.
2. IP-based voice service with full RTP/UDP/IP header. This IP service sends the complete RTP/UDP/IP header over the air interface.
3. IP-based voice service with CRTP header compression. The same IP-based service as in no. 2, above, apart from the introduction of the CRTP header-compression scheme.
4. IP-based voice service with ROCCO header compression. The same IP-based service as in no. 2, above, apart from the introduction of the ROCCO header-compression scheme.

The capacity figures for each of the different cases were derived in a WCDMA system simulator. The simulator models an area with hexagonal cells covered by three sector sites. During the simulation, mobile calls were initiated according to a Poisson process and distributed uniformly over the area. The call duration figures were distributed exponentially, with a mean holding time of 120 seconds. During calls, the mobile terminal generated or received voice frames from a model of the AMR 12.2 codec operating in discontinuous transmission (DTX) mode. Before and after calls, some IP-related signaling (H.323 and RSVP) was transferred over the radio link. The model did not include RTCP and RSVP signaling during calls.

The objective was to estimate capacity at a constant level of quality. However, because it is difficult to measure perceived voice quality in a simulator, the quality measure of a single connection was based on frame error rate (FER). Listening tests indicate that there is a correlation between the perceived voice quality of a connection and FER.

The measure of system quality—to determine system capacity—was not based on

an average connection, but on the number of connections with acceptable quality. In this investigation, adequate system quality was defined as less than 1% FER for at least 95% of all connections. For a given level of system quality, capacity was defined as the maximum load for which this level can be achieved. Capacity was measured relative to the reference case (1) whose capacity was 1.0.

The simulation results (Figure 10) show that for low loads,

- the interference level of the system is moderate; and
- the power control is able to set the individual power levels in such a way that 95% or more of the users have connections with the desired level of quality (FER 1%).

When the load was increased, the interference level also increased, and at some point the system became overloaded. When the system becomes overloaded, interference is excessive and too many connections are bad; that is we reach the capacity limit of the system. When the complete RTP/UDP/IP header was sent over the air interface, capacity dropped to approximately 50% that of the circuit-switched reference case. When CRTP header compression was applied (Figure 10), capacity dropped to approximately 80%; when ROCCO header compression was used, capacity dropped only 10% (to approximately 90%).

The impact on capacity of RSVP and RTCP signaling, which was not included in the simulations, depends on the extent to which it can be reduced and compressed. If the signaling is not reduced, the loss in capacity due to IP-related signaling will equal that caused by compressed headers.

## Conclusion

The widespread growth of the Internet has created a mass market for multimedia and information services. The challenge of providing these services via third-generation wireless systems is twofold: from the market perspective, the challenge is to merge the installed base of users in cellular and Internet environments; and in terms of technology, the challenge is to find common denominators for cellular solutions and efficient Internet access. To succeed in meeting these challenges, third-generation wireless systems must be designed to provide a multitude of services, offering considerable flexibility and cost-effective access with structured quality-of-service handling and en-

suring high radio-spectrum efficiency. The UMTS and GSM/EDGE radio-network architecture and quality-of-service concept are designed to support the needs of present-day and future applications. The concept of bearer services at different network levels makes up the basis for providing end-to-end quality-of-service—transport through the radio-access network is provided via radio-access bearers.

The main objective of the VoIPoW concept is to port voice service to the new packet-data-based platform while maintaining the perceived quality-of-service and spectrum efficiency associated with present-day circuit-switched wireless systems. Given this objective, we see that we cannot choose a single point of implementation. Instead, the main challenge is to find suitable points of implementation that satisfy the voice-service requirements for IP service flexibility or spectrum efficiency. By introducing traffic classification and header compression we can offer a spectrum-efficient VoIPoW service with high voice quality and IP service flexibility.

The aim of designing the third-generation all-IP wireless network is to separate core and radio-access network components, thereby allowing a common packet-switched core network (based on GPRS) to be used for UMTS and GSM/EDGE radio-access networks.

A key objective of third-generation all-IP networks is to provide a capable service platform for IP-based applications. The solutions we have described for audio streams and associated control protocols will advance the UMTS network another step toward becoming a full-fledged service platform that can support demanding services, such as IP-based conversational multimedia.

## REFERENCES

- 1 Nilsson, M.: Third Generation Radio Access Standards. Ericsson Review Vol. 76 (1999):3, pp. 110-121.
- 2 Nilsson, T.: Toward third-generation mobile multimedia communication. Ericsson Review Vol. 76 (1999):3, pp. 122-131.
- 3 Dahlin, S. and Örnulf, E.: Network Evolution the Ericsson Way. Ericsson Review 76 (1999):4, pp.174-181.
- 4 Madfors, M.: Radio Access Network and Quality of Service in Future Wireless Networks. Future Telecom Forum 99, Beijing, China, December 1999.
- 5 Casner, S. and Jacobson, V.: Compressing IP/UDP/RTP Headers for Low-Speed Serial Links, RFC 2508, February 1999.
- 6 Jonsson, L.-E., Degermark, M., Hannu, H. and Svanbro, K.: Robust checksum-based header compression (ROCCO). Internet-Draft (work in progress) draft-jonsson-robust-hc-03.txt. Ericsson Research, January 2000.