PROFESSOR BARRY G. EVANS FREng
DIRECTOR OF CENTRE FOR COMMUNICATION SYSTEMS RESEARCH
AND DEAN OF ENGINEERING, UNIVERSITY OF SURREY
DIRECTOR OF MOBILE VCE LTD

Mobile communications in 2010

VISIONS OF 4G

Image of future phone: Ericsson Ltd

Introduction

Pick up any newspaper today and it is a safe bet that you will find an article somewhere relating to mobile communications. If it is not in the technology section it will almost certainly be in the business section and relate to the increasing share prices of operators or equipment manufacturers, or acquisitions and takeovers thereof. Such is the pervasiveness of mobile communications that it is affecting virtually everyone's life, has become a major political topic and a significant contributor to national GDP.

The major driver to change in the mobile area in the last ten years has been the massive enabling implications of digital technology, both in digital signal processing and in service provision. The equivalent driver now, and in the next five years, will be the all-pervasiveness of software in both networks and terminals. The digital revolution is well underway and we stand at the doorway to the software revolution. Accompanying these changes are the societal developments involving extensions in the use of mobile phones. We began with speech-dominated services but are now experiencing a massive growth in applications involving Short Messaging Services (SMS), together with the beginnings of internet applications using Wireless Application Protocol (WAP) and i-mode (an alternative mobile internet access system developed in Japan). The mobile phone has not simply followed the watch, calculator and organiser as an essential personal accessory – it has subsumed

The digital revolution is well underway and we stand at the doorway to the software revolution

all of them. With the new internet extensions it will also merge the PC, hi-fi and television and provide mobility to facilities previously only available on one network in the so-called Virtual Home Environment (VHE).

> The mobile phone has not simply followed the watch, calculator and organiser as an essential personal accessory – it has subsumed all of them

The digital revolution

The development of mobiles from first-generation analogue systems in 1985 to second-generation (2G) digital GSM in 1992 formed the heart of the digital revolution. But much more than this, the development of GSM was a huge success for standardisation, emanating from Europe and gradually spreading worldwide. (GSM now stands for 'Global System for Mobile Communications'; it originally meant 'Groupe Spécial Mobile'.)

However, worldwide 'roaming' still presents some problems because pockets of US standards IS-95 and IS-136 are still endemic in some countries and are incompatible with GSM. Bandwidth in IS-95 is divided up by Code Division Multiple Access (CDMA) whereas bandwidth division in GSM is achieved by Time Division Multiple Access (TDMA); IS-136 uses a TDMA variant. GSM (2G) is now being extended to allow higher rates of data transmission as well as speech, prior to the introduction of the 3G standard. The extensions (the so-called 2.5G) include GPRS (General Packet Radio Service) and EDGE (Enhanced Data Rates for Global Evolution) as well as WAP and i-mode.

GSM was designed around digital speech services and low bit-rate data that could fit into a speech channel (for example, at 9.6 kbite/s). It is a circuit rather than a packet-oriented network; this is fine for voice transmission but inefficient for data communications. In a circuit-switched network, the user needs to dial in to the network and has exclusive use of a channel (controlled by electronic switches) while information is being transmitted. By contrast any user of a packet-switched network is, in principle, always connected to the network so long as the signal can reach him. Data to be transmitted is broken up into separate, small blocks (packets) and each packet seeks out the most efficient route to the destination address as circuit space becomes available. When received, the packets are reassembled in the correct order to make up the message. In order to address the rapid increase in popularity of internet services, GPRS (General Packet Radio Services) is being added to GSM in 2001 to allow packet (IP – Internet Protocol) communications at up to about 100 kbite/s, over 10 times faster than the average speech channel.

Mobile systems comprise a radio access system together with a supporting core network. In GSM the latter is characterised by the Mobile Applications Protocol (MAP), which provides the key mobility management features of the system.

The third generation

Third generation (3G) mobile phones were standardised in 1999 by the International Telecommunication Union; the standard includes elements from European, US and Japanese protocols. Such 3G systems extend services to include high quality multimedia (multirate) and to use convergent networks of fixed, cellular and satellite components. The radio–air interface standards are based upon WCDMA (Wideband Code Division Multiple Access) which will provide much faster rates of data transmission. The core network has not yet been standardised, but three proposed candidates are: evolved GSM (MAP – Mobile Applications Protocol, giving support for interactive mobile applications), evolved ANSI-41 (from the American National Standards Institute) and an IP-based system. 3G also aims to cater for a diversity of terminal types, including many non-voice terminals such as those embedded in all sorts of consumer products. Bluetooth (another standard not within the 3G orbit, but likely to be associated with it) is a short-range system which addresses such applications. Thus we can envision services from a few bite/s up to 2 Mbite/s.

For broadband indoor wireless communications, standards have emerged to support IP-based services and provide some Quality of Service (QoS) support (that is, to ensure customer satisfaction by providing consistent standards of data delivery even during periods of congestion). Such systems are based on OFDM (Orthogonal Frequency Division Multiplexing – another method for allocating bandwidth) rather than CDMA and are planned to operate in the 5 GHz band.

Whereas 2G operates in the 900 and 1800/1900 MHz frequency bands, 3G is intended to operate in wider bandwidth allocations at 2 GHz. These new frequency bands will provide wider bandwidths for some multimedia services; the first allocations have already occurred in some countries via spectrum auctions (UK, The Netherlands, Germany) or beauty contests (France, Italy). The opportunity has also been taken to increase competition by allowing new operators into the bands as well as extending existing operator licences. These new systems will comprise microcells as well as macrocells in order to deliver the

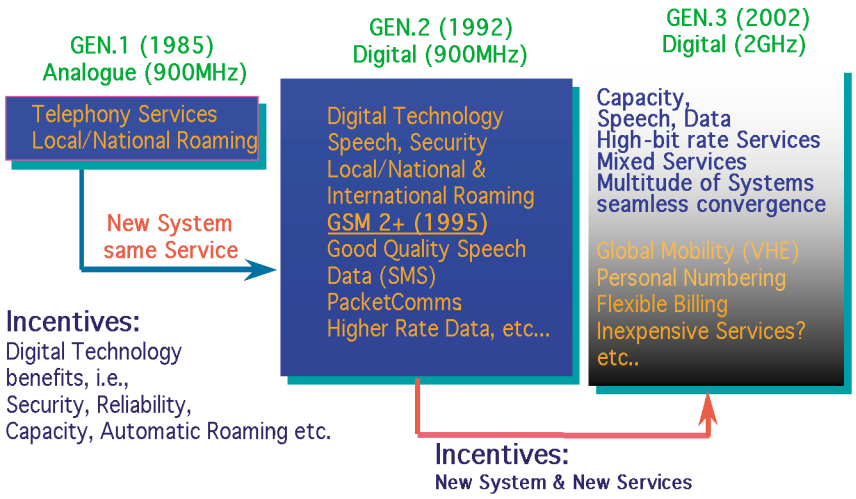
higher capacity services efficiently. 3G and 2G will continue to co-exist for some time with optimisation of service provision between them. Various modes of delivery will be used to improve coverage in urban, suburban and rural areas with satellites (and possibly HAPS – High Altitude Platform Stations, e.g. balloons located at around 20–25 km in the stratosphere) playing a role. The story of the evolution of the mobile radio generations is summed up in the figure below.

As we move from 2G to 3G the convergence of communications and computing is already central to the realisation of the new generation of services and applications. Digital technology enables the dynamic adaptation of systems, and intercommunicating software embedded in networks and terminals allows efficient control of the new networks. This will be further accentuated as we move from 3G to 4G: the range and bit rate of services will be extended and the convergence of fixed, mobile and broadcast networks, service provision and terminal types will come under consideration.

Limitations of 3G and drivers for 4G

From the basic conception of 2G to the time of its roll-out took around ten years and a similar period will apply to 3G, which will commence service in 2001/2002 and reach full deployment by 2005. Thus by 2010 it will be time to deploy 4G networks and, working backwards within the ten-year cycle, the year 2000/2001 is clearly an appropriate time to put forward visions for 4G and a research programme aimed at the key issues. The UK's Mobile Virtual Centre of Excellence (see later) is pursuing a second-phase research programme constructed to meet this aim.

Our starting point has been to look at current trends. Here we see a phenomenal growth in mobile use, with an estimated global user-base that will exceed one billion by 2003. Already mobile communications have exceeded fixed communications in several countries and it is predicted that mobile will subsume fixed by 2010 (that is, fixed–mobile convergence will be complete). Currently short messaging is booming, especially with the younger generation, with averages of over 100 messages per month dominating monthly bills. Business take-up of short messaging via information services is also increasing and providing a start for mobile e-commerce, but is currently very much limited by the available bit rates. This will be improved with the introduction of GPRS (General Packet Radio Service).



Evolution of the mobile radio generations

3G terminal concept
Copyright © Nokia, 2000

At the time of writing we are at the beginning of the stage that follows on from short messaging, which is to provide an efficient and convenient user interface to enable internet content to be accessed on mobile devices – the 'internet to the pocket' era. The European introduction of WAP has been slow to gain market ground but by contrast in Japan, NTT DoCoMo's i-mode system had over 7 million subscribers in May 2000 and is now (December 2000) picking up 50,000 new customers per day. Customers are already browsing the internet, exchanging e-mail, conducting banking and stock transactions, making flight reservations and checking news and weather via HTML-based text

it is predicted that mobile will subsume fixed by 2010

information on their phones. The internet programming language Java is expected to be available on i-mode phones soon, which will allow the downloading of agents, games, etc. and the introduction of location-based services. In Japan, the number of internet phones has now passed the number of wired internet customers and is setting a trend that others will surely follow when 3G opens up more bandwidth and improved quality.

Thus 3G will provide a significant step in the evolution of mobile personal communications. Mobility appears to be one of the fundamental elements in the evolution of the information society. As service provision based on 'network-centric' architectures is gradually giving way to 'edge-centric' architectures, access is needed from more and more places at all times. But can 3G deliver?

It is true that 3G can support multimedia internet-type services at improved speeds and quality compared to 2G. The WCDMA-based radio–air interface has been designed to provide improved high capacity coverage for medium bit rates (384 kbite/s) with limited coverage up to 2 Mbite/s (in indoor environments). The efficiency of packet-mode transmission is also improved by statistical multiplexing on the air, which allocates bandwidth dynamically to allow more channels to be connected and the most efficient use to be made of the available space. However, there are limitations to 3G:

- Extension to higher data rates is difficult with CDMA (Code Division Multiple Access), because of excessive interference between services.
- It is difficult to provide a full range of multirate services, all with different QoS (Quality of Service) and performance requirements, because of the constraints imposed on the core network by the air interference standard (that is, it is not a fully integrated system).

Embedded radio will eventually become as common as embedded microprocessors are today, with perhaps 50 such devices in the typical home

In addition the bandwidth available in the 2 GHz bands allocated for 3G will soon become saturated and there are constraints on the combination of frequency and time division duplex modes of delivery imposed by regulators, to serve different environments efficiently.

Visions of 4G

By the year 2010, one of the key enabling technology developments will be embedded radio – the widespread availability and use of the $1 radio chip, which will evolve from short-range wireless developments such as Bluetooth. Embedded radio will eventually become as common as embedded microprocessors are today, with perhaps 50 such devices in the typical home of which the user is mostly unaware. As they interact – for example, in response to the user arriving home – they will form a Home Area Network (HAN). Similarly such devices will be present in large numbers in vehicles (the Vehicular Area Network, VAN), in personal belongings (the Personal Area Network, PAN), in the public environment, and so on. Such chips will serve as a means of short-range communication between objects and devices, offering capabilities for monitoring and control, in most cases without user knowledge or intervention.

As a person moves between these environments such short-range links will allow his personal profiles and preferences to move with him, with the hotel room automatically configuring itself to his personal preferred temperatures, TV channels/interests, lighting, etc. However, the integration of such links with wide-area mobile access will enable far more powerful service concepts: mobile agents will be able to access this pervasive network of sensors and obtain information on the user's behalf to perform and even pre-empt his needs and wishes.

In the 1G to 2G transition, as well as the transition from analogue to digital, we saw a mono-service to multi-service change. From 2G to 3G, as well as a mono-media to multimedia change, we are also seeing transition from person-to-person to person-to-machine interactions, with users accessing video, internet/intranet and database feeds. The 3G to 4G transition, supported by such technologies, will see a transition towards a predominance of automated and autonomously initiated machine-to-machine interactions.

The 'killer application' for 4G is likely to be the personal mobile assistant (PMA) – in effect the software complement to the personal area network – which will organise, share and (potentially) enhance all of our daily routines and life situations. It will have a range of functions, for example:

the personal mobile assistant (PMA) will organise, share and (potentially) enhance all of our daily routines and life situations

- The ability to learn from experiences and to build on 'personal' experiences – to have intelligence;
- Decision capability to organise routine functions with other PMAs and network databases – for example, diary, travel, holidays; prompts for shopping, haircut, theatre, birthdays, etc.;
- A range of communication modes – voice, image (superimposing images via head-up displays using spectacles or retinal overlays), multiparty meetings including live-action video of the user and his current environment;
- Navigation and positioning information, enabling location-dependent services:
  - Location detection and reporting for children, pets and objects of any sort;
  - Vehicle positioning and route planning, autopilot and pedestrian warnings;
  - Automatic reporting of accidents (to insurance company, rescue services and car dealers);
- Knowledge via intelligent browsing of the internet;
- e-business facilities – purchasing and payment;
- Health monitoring and warnings;
- Infotainment – music, video and maybe virtual reality (VR).

Of course the key to all this is 'mobility' – we need to have the PMA whenever and wherever we are. This inevitably places additional complexity on network and service objects and the agents that process them.

The aim of current research in this area is to develop tools that can be used in 4G software systems. To this end, specific scenarios are being addressed in order to focus the issues:

- e-commerce, including microtransactions, share trading and business internal transactions;
- Home services, ranging from terminal enhancements (e.g. enhancing display capabilities by using a TV screen as the display unit for the terminal) to security systems and housekeeping tasks;
- Transportation systems, in particular itinerary support, ticketing and location services;
- Infotainment on the move, which will require software and terminal reconfiguration and media-adaptation;
- Telemedicine and assistance services: possible scenarios include emergency team support, remote/virtual medical operations and surveillance of heart patients.

4G visions mapping to a research programme

As a result of recommendations from the Foresight Programme's first phase, a UK research venture was formed in 1996 to tackle long-term research into mobile communications and put the UK in the forefront of future radio generation systems. The Mobile Virtual Centre of Excellence (MVCE) is a company limited by guarantee, formed by several leading UK universities and twenty industrial organisations that represent all of the major UK operators and mobile equipment manufacturers.

The company supports a long-range research programme in the participating universities, paid for by industrial contributions (shares in the company) plus government funding, initially from the Foresight Challenge programme and now from EPSRC. The programme of research is set by the company and the work is monitored and steered by the industrial members.

Research is carried out in an entirely new way: academic researchers work alongside industrial colleagues (a true partnership) and collaborate (rather than compete) with colleagues from other universities, working together as a team towards common research aims. The first phase of this research, aimed at 3G to 2000, produced some twenty patents and more than a hundred publications. The current research phase is aimed at 4G and the vision for 2010 is embodied in the five key elements, as shown in the figure on the next page.

Creating a new cadre of researchers

Establishing a major research centre with critical size has been crucial in tackling the wide range of research problems that are needed to make 4G radio systems happen. The following are key issues to be solved:

- Establishing an all-IP network (core + access) with guaranteed quality for a range of services;
- Inter-networking across a 'network of networks' consisting of hierarchical cells all optimised for different environments;
- Integrating broadcast and mobile services;
- Making networks reconfigurable;
- Providing a 'middleware' of reliable and secure software agents;
- Achieving dynamically adaptable wireless access including software terminals and base stations;
- Dynamically allocating bandwidth to achieve spectrum efficiency.

The MVCE has been a major success in establishing the partnerships between academia and industry necessary to tackle these problems. As we now enter a more software-dominated network domain, we are also forming partnerships between the key disciplines of electronic engineering, computer science and applied mathematics that provide the skills to address these problems. In so doing, we hope to produce a new cadre of trained researchers with the necessary skills to flow into industry.

We haven't forgotten the users! The next challenge is to form partnerships with sociologists and psychologists who can perhaps, by working in partnership with engineers and computer scientists, determine future user requirements, questions and constraints, and feed these back to influence the evolution of the technology. ■

Diverse user devices

The user will be served by a wide variety of low-cost mobile devices to access content conveniently and seamlessly. These devices will commonly be wearable – in some cases disposable – and will normally be powered independently of the mains. Devices will interact with users in a multi-sensory manner, encompassing not only speech, hearing and sight, but also the other human senses and biological and environmental data pertinent to the application. Special devices tailored for people with disabilities will be commonplace.

Fully converged services

Personal communications, information systems, broadcast and entertainment will have merged into a seamless pool of content available according to the user's requirement. The user will have access to a very wide range of services and applications, available conveniently, securely and in a manner reflecting the user's personal preferences.

Autonomous networks

Underlying these systems will be highly autonomous adaptive networks capable of self-management of their structure to meet users' changing and evolving demands for both services and capacity. Efficient and cost-effective use of the radio spectrum will be an essential element of their operation, and here too autonomy and self-management will be the norm.

Ubiquitous mobile access

The dominant mode of access to this pool of content will be mobile, accounting for all voice communications, the majority of high-speed information services, and a significant proportion of broadcast and entertainment. Mobile access to commercial and retail services will be the norm, replacing current practices in most cases.

Software dependency

Intelligent Mobile Agents will exist throughout the networks and in user devices, acting continually to simplify tasks and ensure transparency to the user. These Mobile Agents will act at all levels, from managing an individual user's content preferences to organising and reconfiguring major elements of networks.

Key elements of the 4G vision