Broad-Band Wireless Access and Future Communication Networks

REINHARD BECHER, MARKUS DILLINGER, MEMBER, IEEE,
MARTIN HAARDT, SENIOR MEMBER, IEEE, AND WERNER MOHR, SENIOR MEMBER, IEEE

Invited Paper

This paper presents a vision for wireless communication systems beyond the third generation, which comprises a combination of several optimized access systems on a common IP-based medium-access and core network platform. These different access systems will interwork via horizontal and vertical handover, service negotiation, and global roaming. The different access systems are allocated to different cell layers in the sense of hierarchical cells with respect to cell size, coverage, and mobility to provide globally optimized seamless services to all users. This vision requires extensive international research and standardization activities to solve many technical challenges. Key issues are the global interworking of different access systems on a common platform, the implementation of multimode and multiband terminals and base stations by software-defined radio concepts as well as advanced antenna concepts.

Keywords*—*Adaptive antennas, IMT-2000, interworking, MIMO channels, mobile communications, software radio, systems beyond the third generation, technical challenges, UMTS, wireless CORBA.

I. Future Wireless Communication Systems Beyond the Third Generation

According to forecasts, data traffic in wireline networks routed via IP infrastructure surmounts the traditional voice traffic these days. In addition to this, IP-based data traffic doubles every four months. This process is significantly driven by the deregulation and cost reduction for data transmission (see Fig. 1). This change of paradigm (from voice to data as fastest growing market) has significant impact to the further development of the telecommunication market [1].

In mobile communications, today's main traffic type in terms of users is still voice. However, the usage of the short message service (SMS) in second-generation systems like GSM has increased significantly during the past years. In countries with a high mobile penetration, the SMS generates already over 10% of the operators' revenues. The UMTS Forum expects that mobile multimedia services will create around 60% of the whole traffic in Europe (in terms of transmitted bits) by 2010 (see Fig. 2). With the deployment of wireless application protocol (WAP) [3] supporting terminals, a first step toward "Internet on Air," the availability

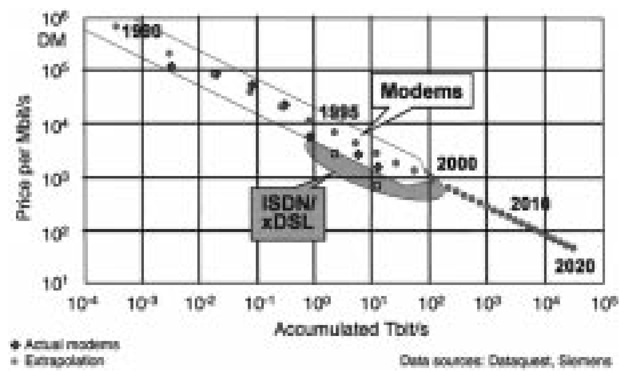

Fig. 1. Digital access technologies—network termination experience curve [2].

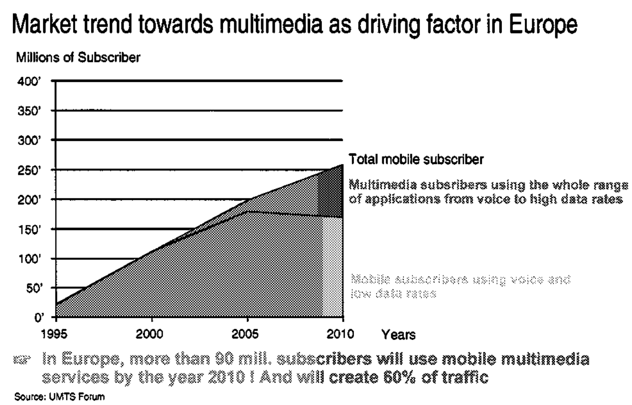

Fig. 2. Use of mobile multimedia services in Europe by the year 2010 [4].

Manuscript received March 13, 2000; revised September 12, 2000.
The authors are with Siemens AG, ICM CA, Information and Communication Mobile, D-81359 Munich, Germany (e-mail: Reinhard.Becher@icn.siemens.de; Markus.Dillinger@icn.siemens.de; Martin.Haardt@icn.siemens.de; Werner.Mohr@icn.siemens.de).
Publisher Item Identifier S 0018-9219(01)00452-2.

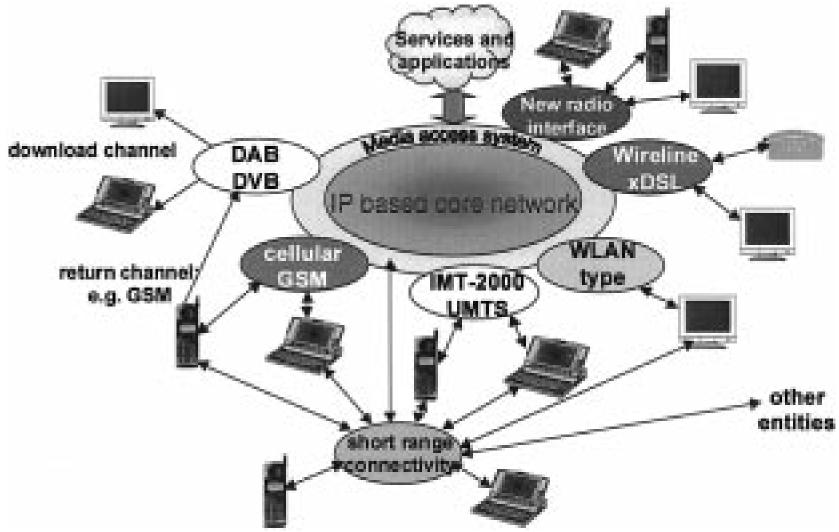

Fig. 3. Seamless future network including a variety of access technologies.

and usage of Internet like mobile data applications will increase significantly. Whereas the first WAP supporting terminals uses circuit switched data bearer, the next generation will use packet data networks, e.g., GPRS for second-generation systems and IMT-2000/UMTS for third-generation systems. Currently these systems are in the deployment or planning phase.

Future communication systems beyond the third generation will be characterized by a horizontal communication model between different access technologies such as cellular, cordless, wireless local area network (WLAN)-type systems, short-range connectivity, broadcast systems, and wired systems as, for instance, explained in [5]. They will be integrated on a common platform to complement each other in an optimum way and to satisfy different service requirements in a variety of radio environments. These access systems will be connected to a common, flexible, and seamless IP core network.

Users will have a single number for all access technologies. A new media access system (generalized access network) that contains the mobility management connects the core network to the appropriate access technology. Global roaming is required for all access technologies. Key requirements include the interworking between these different access systems in terms of horizontal (intrasystem) and vertical (intersystem) handover as well as seamless services with service negotiations including mobility, security, and quality of service (QoS). They will be handled in the (newly developed) media access system and in the core network. Fig. 3 shows this vision of a seamless network that includes a variety of interworking access systems, which are connected to a common IP-based core network. The media access system connects each access system to the common core network.

The different access systems are organized in a layered structure, which can be compared to hierarchical cell structures in cellular mobile radio systems. This concept facilitates an optimum system design for different application areas, cell ranges, and radio environments, since a variety of access technologies complement each other on a common platform. This layered structure is illustrated in Fig. 4.

- *Distribution layer:* The distribution layer contains emerging digital broadcasting (or distribution) systems such as digital audio broadcasting (DAB), digital video broadcasting (DVB) [6], and satellite systems that have a global coverage and support large cells, full mobility, as well as global access. Individual links are not necessarily needed for broadcasting services. But this technology can also be used as a broad-band downlink channel to provide, for instance, fast Internet content. Basically, all other access systems may be used as return channels for data requests and acknowledgment signaling in highly asymmetric services.
- *Cellular layer:* The cellular layer provides a high-system capacity in terms of users and data rates per unit area. It will consist of second-generation (e.g., GSM and its evolution) and third-generation mobile radio systems (IMT-2000/UMTS: UTRA FDD and TDD) for data rates up to 2 Mb/s [7]. This layer is characterized by full coverage, full mobility, and global roaming and is well suited for multimedia applications. The cellular layer is designed to support individual links.
- *Hot spot layer:* The hot spot layer also supports individual links and is intended for very high data rate applications. It should be employed in hot spots such as in company campus areas, conference centers, and airports. The hot spot layer contains WLAN-type systems as HiperLAN Type 2, IEEE 802.11, or mobile multimedia advanced communications (MMACs). These systems are flexible with respect to asymmetric data services with respect to uplink and downlink, supported data rates, and adaptive modulation. In contrast to cellular systems, this layer contains systems that are characterized by a shorter range and mainly local coverage with local mobility to facilitate an economic system deployment. Global roaming is possible and will be required. Full coverage, however, is not expected.

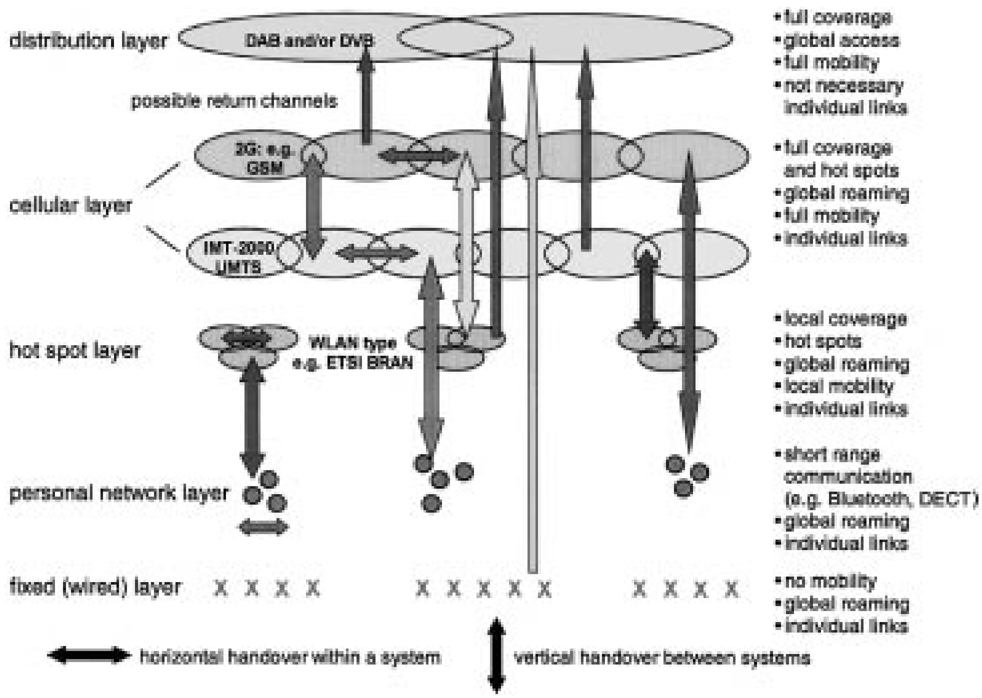

Fig. 4. Layered structure of a seamless future network.

- *Personal network layer:* The personal network layer will mainly be used in office and home (household) environments. Different equipment (laptops, printers, personal digital assistants, etc.) and appliances (refrigerators, toasters, washing machines, smart sensors, etc.) can be connected to each other to provide short-range connectivity via systems such as Bluetooth and DECT. These systems can also be used to connect the equipment directly to the medium-access system or to multimode terminals that can transmit on one of the other network layers and are, of course, also equipped with a short-range connectivity system. This facilitates an efficient interconnection between the devices as well as a connection from these devices to the public network. The systems of the personal network layer in general do not support mobility. Global roaming, however, should be ensured.
- *Fixed (wired) layer:* The fixed layer contains fixed access systems such as optical fiber (e.g., FTTx), twisted pair systems (e.g., xDSL), and coaxial systems (e.g., CATV). Furthermore, fixed wireless access or wireless local loops can be included in this category. Fixed access systems should support mobility (portability) in the sense of terminal portability such that global roaming is feasible. These systems of the fixed layer have a high capacity and in general do support individual links.

The network ensures the interworking between these systems on the common platform by horizontal handover within an access system and, in particular, by vertical handover between different access systems. In general, vertical handover takes place between different layers of the common platform. Vertical handover is combined with service negotiations to ensure seamless service, because in general different access systems support different user data rates and other bearer and service parameters. The interworking, mobility management, and roaming will be handled via the medium-access system and the IP-based core network (Fig. 3).

Multimode terminals and new appliances are key components of such a seamless network. Among other features, these terminals might comprise a camera, a screen for video and high-resolution Internet applications as well as transceivers that provide short-range connectivity for *ad hoc* networking with other devices. Such terminals may be adaptive with respect to different standards and channel conditions due to sophisticated signal processing algorithms.

Future wireless systems beyond the third generation require the integration of various network technologies stretching from current computer networks to the wireless world. Different enabling technologies starting from low-level signal processing up to the management of distributed network entities must be considered when heterogeneous networks merge together.

In this technology survey, the requirements and possible solutions for both are derived. This vision will be established by allowing an optimized usage of the spectrum resources, a provision of increased flexibility on the terminal side and in the network, the support of global roaming, the personalization of services as well as self-planning and self-configuration of networks.

A variety of networks has created an important need to enhance current middleware concepts for the wireless world. In particular, the reconfiguration of terminals and other network components demands a network management vision with a seamless connection between heterogeneous networks and a new transparency throughout. The forthcoming standardization of wireless CORBA [8] may be a key driver in the integration of existing and future networks by introducing reliable reconfiguration among the network entities and terminals. A framework for reconfiguration is presented with a proposed terminal architecture, and it is shown how the current well-known mobile execution concepts based on the virtual machine (VM) approach can be combined with the CORBA methodology for a reliable reconfiguration. The reconfiguration process can be seen as a special user application and the individual phases in the reconfiguration process are outlined, which need to be taken into account in the forthcoming standardization of wireless CORBA.

The remainder of this paper is organized as follows. The technical challenges for future wireless communication systems beyond the third generation are outlined in Section II. They include the medium-access system illustrated in Fig. 3 and new radio technologies. Then, capacity enhancing features are described in Section III. These features include smart antenna concepts (Section III-A), uplink processing with smart antennas (Section III-B), downlink processing with smart antennas (Section III-C), and multiple antennas at the base station and at the terminal (Section III-D). Finally, Section IV deals with software radio aspects.

II. Technical Challenges

Many technical challenges have to be solved by extensive research efforts to realize the vision for wireless communication systems beyond the third generation described in the previous section. Key concepts include the interworking between different access systems on a common platform as well as the required adaptive multimode and multiband terminals for different access systems and a wide range of services [5]. There are challenges in several areas such as the radio interface, the radio access network and the core network, implementation issues, and services-related issues. New technologies for tackling increased flexibility and improved transmission performance have to be developed. Furthermore, there are demands on the increase of the transmission capability and versatility of networks and terminals through the support of reconfiguration, the provision of self-planning and self-configuration of networks as well as the support of personal services and service mobility. Future system concepts have to address the terminal architecture in addition to concepts that enable the reliable reconfiguration across network entities. Enabling technologies include analog RF and analog and/or digital IF. New designs evolve to realize linear power amplifiers with a high dynamic range, enhanced digital signal processing, and software techniques that facilitate the programming of distributed components [9].

From the user's point of view, the man–machine interface has to be easy to use and self-explanatory to enable people—even handicapped and elderly people—to use easily advanced services. The content has to be adapted automatically to the actual bearer capability of the used access system.

To this end, extensive international research activities are necessary to solve technical issues and to prepare the consensus building for the international standardization of new ideas and concepts. Major challenges include the interworking of systems via vertical handover, global roaming, and the optimization of evolving and emerging access systems as well as the radio access and core network.

A. Medium-Access System

Higher layer protocols in the access network (medium-access system) are an important area to improve the system performance even further. These are, for example:

- self-optimizing networks and automatic planning;
- resource allocation algorithms with respect to varying traffic load, services, bearer capabilities, radio environment, and channel conditions;
- dynamic frequency allocation;
- interworking of different access systems on higher layers via horizontal and vertical handover and service negotiation;
- network management.

The interworking of different access systems on an IP-based common platform via the medium-access system with horizontal and vertical handover results in several challenges on the core and the medium-access system. Major topics are related to the improvement and extension of IP for mobile applications and the optimization for radio transmission, such as:

- support of real-time and nonreal-time services with respect to service requirements as QoS and especially delay requirements for real-time services;
- service negotiation for seamless services versus the available access systems and bearer capabilities;
- mobility management including handover and roaming;
- security mechanism such as authentication, authorization, and accounting.

B. New Radio Technologies

In spite of the fact that only existing and emerging communication systems have been mentioned in Section I, systems that use new radio technologies will also be integrated into such a seamless future network. Such new radio technologies include, for instance, generalized multicarrier (GMC) CDMA systems that are capable of multiuser interference (MUI) elimination and intersymbol interference (ISI) suppression, irrespective of the encountered wireless frequency selective channel. This channel-independent feature can be obtained via special block-precoding schemes that are described in [10].

There will also be layered *self-organizing networks*, where a whole service area is divided into multiple subservice areas. Only one base station serves one subservice area, as depicted in Fig. 5. Mobile stations that are on the outside of the coverage area of a base station are connected to one of the adjacent base stations through several repeating mobile stations in a self-organizing fashion (multihop architecture). Thereby, a layered self-organizing wireless network can cover a wide area with fewer conventional (fixed) base stations than a conventional cellular network if a sufficient

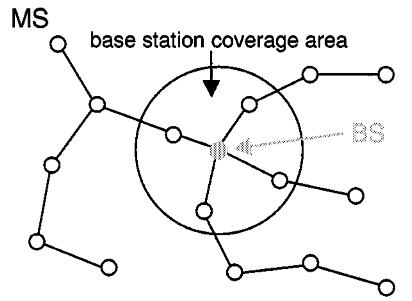

Fig. 5. Subservice area of a layered self-organizing network [11].

penetration of mobiles is available [11]. Hence, less infrastructure is required. Note that the cell size of a cellular system decreases with an increasing demand for capacity. In this case, network planning is not really feasible anymore. Therefore, sophisticated network planning is not necessary and shadowing can be overcome efficiently. Compared to a conventional cellular system, such a self-organizing wireless network requires smaller transmission powers that lead to a reduced overall interference situation in the network.

Mobile terminals can, however, only be used as relays in a particular area if a sufficient subscriber density is available. Moreover, the capacity of such terminals must be large enough.

III. Summary of Capacity Enhancing Features

Future systems have to use the frequency resources as efficient as possible. Therefore, several physical layer related techniques have to be investigated. These include:

- optimization of evolving and emerging access systems by improved modulation and channel-coding schemes for a further enhancement of the spectrum efficiency and system performance;
- advanced detection schemes such as multiuser detection and interference cancellation to gain from the *a priori* knowledge about intra- and intercell interference signals;
- compression techniques for source coding to reduce the needed user data rate.

All the signal processing algorithms that are implemented in practice realize a tradeoff between the achieved performance gain and the involved computing complexity. Improved algorithms to support these mainly physical layer issues are as follows.

- *Link Adaptation* according to the channel conditions and services for a better usage of the frequency resources and an improved system performance. Such link adaptation techniques include adaptive modulation, adaptive equalization, adaptive power control, and adaptive channel coding. Note that a feedback channel from the terminal to the base station might be required to insure a robust performance. If the channel is fading, adaptive modulation techniques can achieve a significant fraction of the channel capacity [12]. For multiple-input–multiple-output (MIMO) systems this implies varying the modulation format and bit rate across antennas and across time. Alternative modulation independent signal processing schemes for MIMO channels are discussed in Section III-D.
- *Adaptive Radio Resource Management* to take into account the present traffic load. This includes dynamic radio channel allocation, power control techniques, admission- and load control strategies, congestion control, handover strategies, and resource scheduling. The cell size, their layout, and hierarchical cell structures must also be considered.
- *Adaptive Spectrum Management* techniques between different access systems and the investigation of *coexistence conditions* between different radio access systems.
- *Advanced Antenna Concepts* to improve the link quality and channel capacity (Fig. 6).

These concepts are used to increase the channel capacity of the radio link. Diversity concepts like frequency diversity, time diversity, polarization diversity, or space diversity reduce the impact of fading due to multipath propagation. As an example for capacity enhancing features, an overview of adaptive antenna concepts is given in the sequel. These are key concepts to use the scarce frequency spectrum as efficiently as possible without major impacts on evolving access systems. Third-generation mobile radio systems (IMT-2000/UMTS) have already taken into account necessary prerequisites for adaptive antenna concepts. However, an economic implementation of the different RF front ends and the baseband signal processing are technical challenges.

A. Smart Antenna Concepts

Smart antennas can be realized by an antenna array and intelligent (baseband) signal processing. The antenna array may be located at the base station and/or the mobile terminal depending on the array size. First, we assume that only the base station is equipped with an antenna array. The antenna array is used for adaptive directional reception (on the uplink) and adaptive directional transmission (on the downlink). Thereby, an increased antenna gain and an increased diversity gain are realized toward the desired user, as illustrated in Fig. 7. At the same time, less interference is received from the other directions on the uplink or transmitted in the other directions on the downlink. Therefore, more users can be accommodated by the system and a corresponding capacity increase is achieved.

In contrast to user-dedicated traffic channels, the broadcast channels should be transmitted in an omnidirectional or sectorized fashion. Hence, there is neither an additional antenna array gain nor a diversity gain for these broadcast channels. If no countermeasures are taken, the broadcast channels have a smaller maximum range than the user-specific channels that can be transmitted in a space-selective fashion if the transmit power is the same for both types of channels. To compensate this effect, space–time coding [13] and/or a higher spreading gain may be used for the omnidirectional or sectorized broadcast channels.

1) Smart Antennas in TDMA-Based Cellular Systems: In TDMA-based cellular systems, smart antenna concepts can

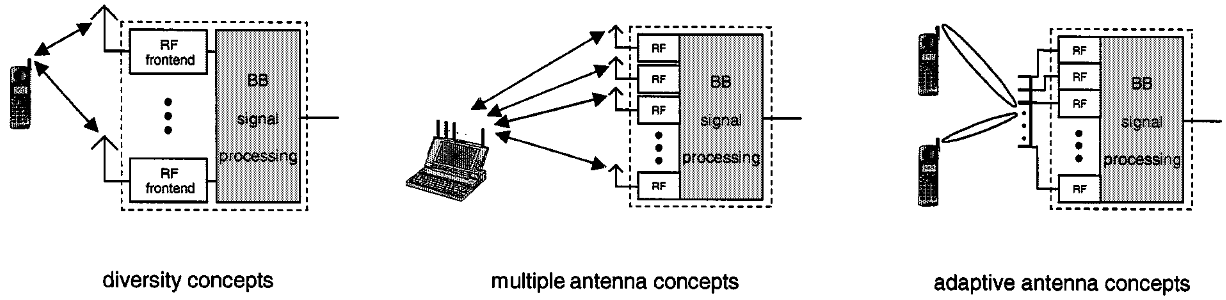

Fig. 6. Advanced antenna concepts.

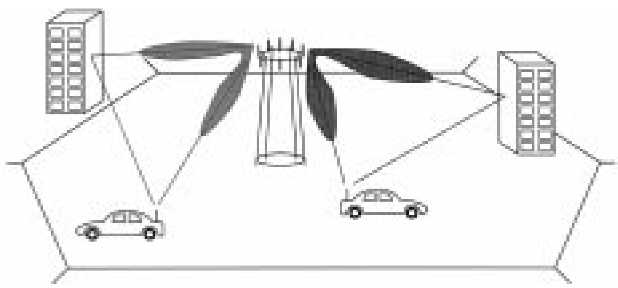

Fig. 7. Smart antennas at the base station of a mobile radio system illustrated for two cochannel users.

be classified according to their area of application [14] as described in the following.

a) High-sensitivity reception (HSR): In the first case, smart antennas are only used on the uplink to increase the uplink antenna gain and suppress cochannel interference. Thereby, the system capacity is increased on the uplink, and the transmit power of mobile stations in a cell served by smart antennas can be reduced. On the more critical downlink, the system capacity is not increased. A range extension and a reduction of base station sites is only possible by balancing the downlink with an increased transmit power.

b) Spatial filtering for interference reduction (SFIR): In the SFIR concept, smart antennas are employed on the uplink as well as the downlink (balanced operation). Thereby, an increased antenna gain and a significant interference reduction can be achieved on both links. Therefore, the frequency reuse distance can be reduced, which leads to an increase of capacity due to a tighter frequency reuse and an increased load in the network. Note that only one cochannel user is served per time and frequency slot.

c) Space-division multiple access (SDMA): In this most advanced mode of operation, which includes uplink and downlink processing, *several cochannel users* can be served on a single time and frequency slot. Spatial multiplexing through a directional reuse of the spectrum resources leads to a direct capacity increase.

2) Smart Antennas in CDMA-Based Cellular Systems: In CDMA-based systems, there are *several cochannel users* that are inherently separated by different spreading codes. More and more cochannel users can enter the system until a certain multiple-access interference level has been reached. Here, directional transmission and reception at the base station via smart antennas causes a significant reduction of the multiple-access interference. Therefore, more users can be accommodated by the system and a corresponding capacity increase is achieved. Similar concepts as for TDMA-based systems are applicable.

B. Uplink Processing with Smart Antennas

1) Space-Only Processing: In case of several antennas at the base station, space-only processing (spatial beamforming) is sufficient if the wireless channel has a small delay spread with respect to the chip or symbol duration. Otherwise, a subsequent temporal equalizer is necessary, yielding a hybrid structure that consists of two stages and performs decoupled spatial and temporal processing. Spatial beamforming techniques can be classified into nonadaptive and adaptive techniques.

a) Switched beams: In the switched beam approach, the direction of the beams is not adaptive. There is a beamformer in the RF stage that forms multiple nonadaptive beams. These fixed beams may be generated by independent fixed antennas or by an RF beamforming network that performs a DFT (Butler matrix). In practice, there might be 4–8 beams per sector (assuming that a base station has three sectors of 120°). A switch is used to select the best or the best two beams for the receiver and the transmitter. Existing diversity antennas can easily be replaced by a switched beam antenna system.

The performance gains are due to the array gain, the diversity gain, and a reduced interference if the desired signal and the interference are well separated in angle such that they fall into different beams.

b) Adaptive beamforming: Adaptive spatial beamforming techniques (space-only processing) can be based on the estimated spatial signature, which might be a superposition of several array steering vectors, and/or the transmitted training symbols as, for instance, in the spatial LMS algorithm. They are more flexible than the fixed beam approach, since they provide an improved suppression of interferers by spatial notching and an increased gain for the desired signal. Digital beamforming is achieved via baseband signal processing. In general, M transceivers are required for M antennas.

2) Space–Time Processing: Joint space–time processing schemes achieve a better performance than purely spatial beamforming algorithms. An efficient implementation of joint space–time processing for joint (multiuser) detection in UTRA TDD is described in [15]. Here, the intracell interference is eliminated via joint detection in the space–time

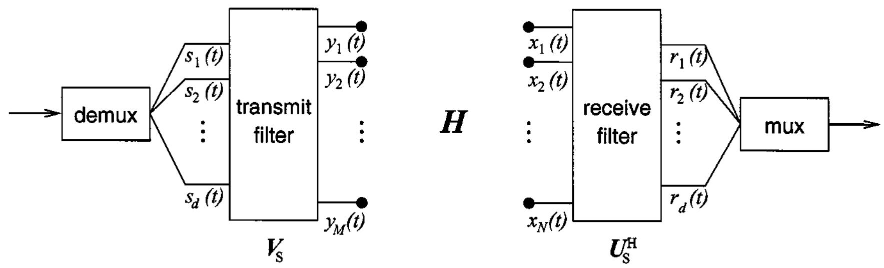

Fig. 8. Illustration of a MIMO wireless communication system with M transmit and N receive antennas, where d parallel data streams are transmitted via d virtual subchannels ($d \leq \min\{M, N\}$).

domain. Moreover, the fact that the base stations of UTRA TDD are frame-synchronized facilitates an efficient suppression of dominant cochannel interferers from adjacent cells (intercell interference) via smart antennas [16]. Several space–time Rake receiver structures for UTRA FDD are presented and compared in [17].

C. Downlink Processing with Smart Antennas

Due to the fact that the uplink and the downlink operate on the same frequency in TDD systems, parameters (e.g., spatial covariance matrices) estimated on the uplink can also be used to calculate the weights for *downlink* beamforming techniques. An efficient spatial downlink beamforming scheme for UTRA TDD is presented in [18]. In FDD systems, the average uplink directions of arrival at the base station are also equal to the average downlink directions of departure from the antenna array at the base station. Therefore, the spatial weights for space-selective downlink transmission could be obtained by transforming the spatial weights estimated on the uplink according to the frequency gap between the uplink and the downlink direction [19].

D. Multiple Antennas at the Base Station and at the Terminal

The throughput can be increased even further if an antenna array is used not only at the base station, but also at the terminal. Multiple antenna concepts are a further extension of diversity concepts. They exploit independent multipath transmissions between the different antenna elements on the base station and the terminal side. As an example, Fig. 8 depicts a MIMO wireless communication system with M antennas at the transmitter and N antennas at the receiver. Such MIMO wireless channels have a significantly higher capacity than single-input–single-output (SISO) wireless channels. They are particularly interesting for laptop type terminals that require high data rates and operate at higher carrier frequencies, e.g., in the 5-GHz range. Multiple antennas can easily be installed on such a terminal due to the fact that the antenna array requires a smaller physical size (aperture) for higher frequencies.

The basic idea is to reuse the same frequency band several times by creating several virtual and uncorrelated subchannels such that several different data streams can be transmitted in parallel. These subchannels experience independent short-term fading. To illustrate this concept, assume that the wireless channel is frequency nonselective and that h_{ij} is the propagation coefficient from the jth transmit antenna ($j = 1, 2, \ldots M$) to the ith receive antenna ($i = 1, 2, \ldots N$). The channel encountered by a subcarrier of an OFDM systems can, for instance, be modeled as frequency nonselective if the subcarrier is sufficiently narrowband. In practice, the receiver estimates the propagation coefficients by exploiting the knowledge of pilot tones and/or pilot symbols inserted in the transmitted data vectors. A blind technique to estimate the propagation coefficients has been derived in [20]. In a TDD system, the propagation coefficients in both directions (uplink and downlink) can be considered as equal if the time between uplink and downlink transmission is smaller than the coherence time of the channel.

In the sequel, we give a simple illustrative example of how several virtual and uncorrelated subchannels can be created in a MIMO wireless communication system. Let $\boldsymbol{H}$ be the propagation matrix of size $N \times M$ that contains all propagation coefficients h_{ij}. Moreover, the vector

$$\boldsymbol{x}(t) = \boldsymbol{H}\boldsymbol{y}(t) + \boldsymbol{n}(t) \in \mathbb{C}^N \tag{1}$$

comprises the array measurements at the receiving antenna array as depicted in Fig. 8, where the vector $\boldsymbol{n}(t) \in \mathbb{C}^N$ contains the additive noise. Here, it is assumed that the components of the noise vector are uncorrelated zero mean random variables with variance $\boldsymbol{\sigma}^2$ such that

$$E\left\{\boldsymbol{n}(t)\boldsymbol{n}^H(t)\right\} = \boldsymbol{\sigma}^2\boldsymbol{I}$$

where the superscript ^{H} denotes complex conjugate transpose. Moreover assume that $E\{\boldsymbol{s}(t)\boldsymbol{s}^H(t)\} = \boldsymbol{I}$ and $E\{\boldsymbol{s}(t)\boldsymbol{n}^H(t)\} = \boldsymbol{0}$. If there is only a line of sight (and no reflections) between the transmitter and the receiver, the propagation matrix $\boldsymbol{H}$ has rank one, since it can be decomposed as

$$\boldsymbol{H} = \alpha\boldsymbol{a}_N(\phi)\boldsymbol{a}_M{}^T(\theta) \in \mathbb{C}^{N \times M} \tag{2}$$

where $\boldsymbol{a}_M(\theta) \in \mathbb{C}^M$ and $\boldsymbol{a}_N(\phi) \in \mathbb{C}^N$ are the array steering vectors of the transmit and the receive array, θ is the corresponding direction of departure at the transmit array, ϕ is the corresponding direction of arrival at the receive array, and α denotes the path loss. In this case, there is only one (line-of-sight) connection between the base station and the terminal, and no virtual subchannels can be created. In this case, the received signals at all antennas experience the same fading. On the other hand, if there is multipath propagation and/or scattering in the vicinity of the terminal and/or the base station, the propagation matrix $\boldsymbol{H}$ has a larger rank [21]. If the angular spread is very large, $\boldsymbol{H}$ might even have full rank, i.e.,

$$\operatorname{rank} \boldsymbol{H} = \min\{M, N\}.$$

If d is the numerical rank of the estimated propagation matrix $\boldsymbol{H}$, d virtual subchannels can be created. The parameter d can, for instance, be obtained from a channel estimate via the Akaike information criterion (AIC) or the maximum description length (MDL) criterion [22]. Let us denote the singular value decomposition (SVD) of $\boldsymbol{H}$ as

$$\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H = [\boldsymbol{U}_S \quad \boldsymbol{U}_O] \begin{bmatrix} \boldsymbol{\Sigma}_S & \\ & \boldsymbol{\Sigma}_O \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_S^H \\ \boldsymbol{V}_O^H \end{bmatrix} \quad (3)$$

where the unitary matrices $\boldsymbol{U} \in \mathbb{C}^{N \times N}$ and $\boldsymbol{V} \in \mathbb{C}^{M \times M}$ contain the left and right singular vectors, respectively, the matrix $\Sigma \in \mathbb{R}^{N \times M}$ contains the singular values of $\boldsymbol{H}$, the block matrix $\Sigma_S \in \mathbb{R}^{d \times d}$ is a diagonal matrix that has the d largest singular values of $\boldsymbol{H}$ on its main diagonal and zeros elsewhere, and the block matrices $\boldsymbol{U}_S \in \mathbb{C}^{N \times d}$ and $\boldsymbol{V}_S \in \mathbb{C}^{M \times d}$ consist of the left and right singular vectors that correspond to the largest d singular values of $\boldsymbol{H}$, respectively.

At the transmitter, the bit stream is coded and modulated, demultiplexed into d subpackets, multiplied by the transmission matrix $\boldsymbol{V}_S$, and launched into the channel using M transmit antennas, as illustrated in Fig. 8. After the transmission over the wireless channel, the signals are received by N antennas at the receiver and processed to recover the original bit stream. Assume that the column vector $\boldsymbol{s}(t)$ comprises the d data symbols $s_k(t)$ $(k = 1, 2, \cdots, d)$ that are transmitted in parallel at time t. The optimum power allocation policy for the d parallel subchannels that satisfies the minimum mean square error (MMSE) criterion under a total average transmitter power constraint has been derived in [23]. Then, the components of the vector $\boldsymbol{y}(t) = \boldsymbol{V}_S\boldsymbol{s}(t) \in \mathbb{C}^M$ are transmitted by the M transmit antennas, as depicted in Fig. 8. In the receiver, the receive array measurements $\boldsymbol{x}(t)$ in (1) are multiplied by $\boldsymbol{U}_S^H$ to yield

$$\boldsymbol{r}(t) = \boldsymbol{U}_S{}^H\boldsymbol{x}(t) = \boldsymbol{U}_S{}^H\boldsymbol{H}\boldsymbol{V}_S\,\boldsymbol{s}(t) + \boldsymbol{U}_S{}^H\boldsymbol{n}(t) \in \mathbb{C}^N \quad (4)$$

as illustrated in Fig. 8. By inserting the SVD of $\boldsymbol{H}$ in (3) into (4), we find that

$$\boldsymbol{r}(t) = \boldsymbol{\Sigma}_S\boldsymbol{s}(t) + \boldsymbol{n}'(t) \in \mathbb{C}^N \quad (5)$$

where the transformed interference-plus-noise vector equals $\boldsymbol{n}'(t) = \boldsymbol{U}_S^H\boldsymbol{n}(t)$. Due to the unitary nature of this transformation, the components of $\boldsymbol{n}'(t)$ are also uncorrelated zero mean random variables.

Hence, the original channel $\boldsymbol{H}$ between $\boldsymbol{y}(t)$ and $\boldsymbol{x}(t)$ has been transformed into d independent virtual channels between $\boldsymbol{s}(t)$ and $\boldsymbol{r}(t)$. These are represented by the diagonal elements of $\boldsymbol{\Sigma}_S$. Note that these d virtual channels should experience independent fading since they are due to independent propagation paths. Hence, we get a d-fold diversity gain on top of the antenna array gains of the transmit array and the receive array.

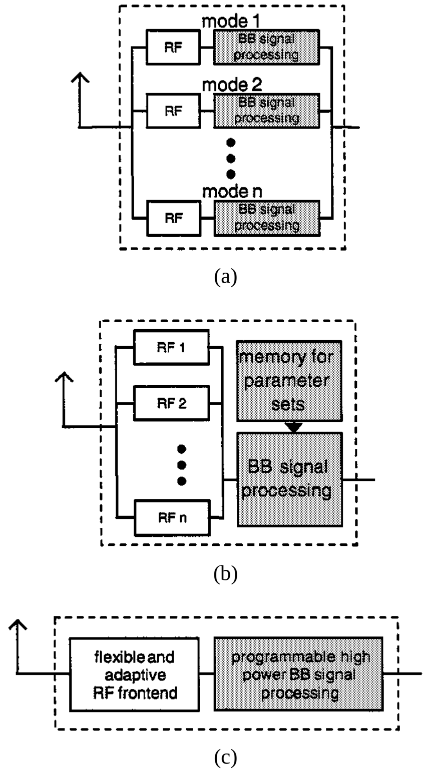

Fig. 9. (a) Illustration of multimode terminals with parallel modes, (b) multimode terminals with software-defined signal processing, and (c) fully adaptive software terminals.

Due to the high capacity of the MIMO wireless channel [21], the described transceiver structure that uses multiple antennas at the receiver and the transmitter is a promising concept to transmit high data rates in wireless communications. Such concepts have to be evaluated under realistic channel conditions by taking into account fading and estimated channels impulse response matrices.

IV. Software Radio

The seamless future network in Fig. 3 comprises several access systems with seamless interworking for different applications and radio environments. Users will only gain from this flexible concept if multimode and multiband terminals with low power consumption and reasonable size are feasible. A variety of terminal types such as PDAs, notebooks, and handsets support these applications. Other technical challenges include improved display techniques and battery technology. Several concepts for the terminal implementation are under discussion (Fig. 9). Note that these techniques will also be applied to base station equipment. In the simplest concept, several fixed modes are implemented in *parallel*. However, this straightforward concept is inflexible for future improvements.

There are several terms often used in the context of software radio and, in principle, two concept ideas can be outlined. In the *software-defined radio* concept, many or all of the blocks already implemented in hardware (or maybe in

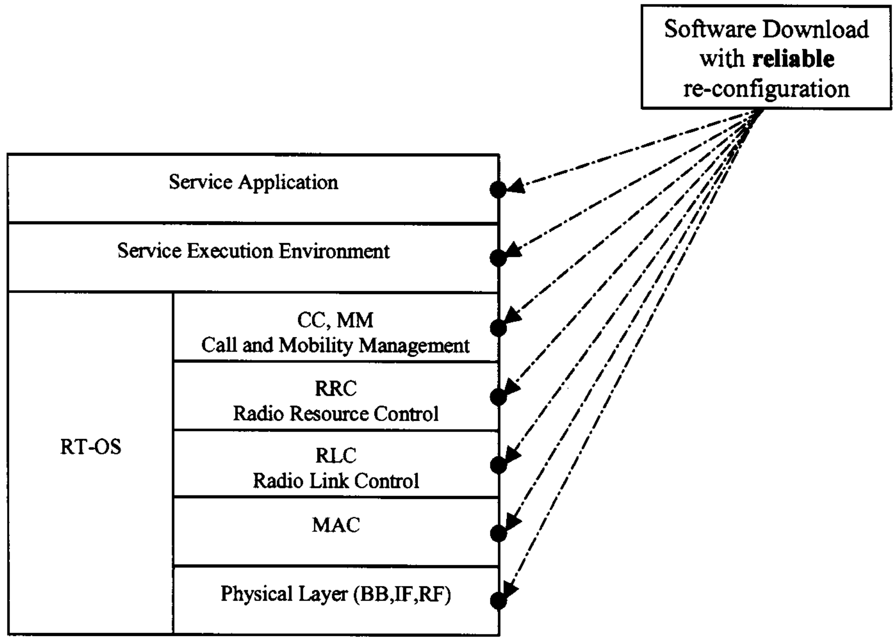

Fig. 10. Reliable software download for the terminal.

software) must be configured by setting certain parameters or by downloading parameter sets of different access systems to the signal processing unit (block diagram in the middle of Fig. 9). The architecture of a software-defined radio needs features outside of the regular design.

The *software radio* shown in the right block diagram of Fig. 9 is the most flexible configuration [24]. Here, all signal processing functions are implemented via programmable digital signal processors (DSPs). It has a great flexibility such that it can be programmed for emerging standards or updated dynamically with new software without any changes in the hardware. The software radio will become feasible due to the progress in semiconductor technology and the increase in available signal processing power.

Concepts of software download at different levels of the protocol stack from applications down to the physical layer have major potential implications not only for terminals and base stations, but also for the whole network. This will determine how applications and services are delivered and in the longer term will influence spectrum usage and applications. A key element of the download is to ensure a reliable reconfiguration over the disturbed air interface. Therefore, research work has to concentrate on this facet. This is illustrated in Fig. 10.

A. Distributed Objects in Future Wireless Networks

The system concepts for mobile communication in the second generation have their main focus on the support of voice communication, which is currently the main traffic type in mobile communications. The development and support of circuit-switched data services in the second-generation mobile networks could be regarded as an addition, as it was in the public wireline networks. Third-generation systems are focused on data services with packet-oriented transmission.

With the advent of computer and data networks (intranet, Internet, extranet) and their rapid growth in the past decade, an alternative approach to the networkcentric application provisioning is required. This approach, known as client–server architecture, differentiates not only by the used protocols, but also by the architectural approach. A basis for the success of the introduction of this architectural type was the development of object-oriented software techniques for distributed systems like COM/DCOM (distributed component object model), JAVA RMI (remote method invocation), CORBA (common object request broker architecture), and mobile agents. Besides the techniques and protocols used to enable web browsing [hyper transfer text protocol (HTTP) and hyper text markup language (HTML, XML)] and transmission protocols (e.g., TCP, UDP, RTP, RSVP, IP, and HTTP), the mentioned technologies are designed to enable the support of distributed systems. But other aspects, such as multithreads and persistent storing, need to be addressed to complete the requirements of new architectures for distributed systems in mobile radio networks.

Another key issue that shall be covered is the increasing demand for open and secure interfaces for the provision of new services and applications to the customer. These new technologies for data networks are mostly driven by computer companies and their fora (e.g., IETF, W3C, OMG, ...). Due to this development, the traditional standardization bodies for telecommunication (e.g., ITU, ETSI, ...) lose some of their influence.

Several new groups, e.g., 3G.IP, MWIF, have started to work on this topic. The question is which technologies will dominate the "mobile Internet." Computer technologies will have a significant impact—they will presumably dominate the "mobile Internet." The reasons for that assumption are as follows. First, with the provision of higher data rates over the radio link, the subscriber expects the same content in the same or a similar way [same look and feel, virtual home environment (VHE)] as being online over wireline. Therefore, the Internet network technologies will extend to the radio access. Another reason is that today's telecommunications infrastructure is (mostly) a monolithic systems. Proprietary software runs on proprietary hardware. By contrast, computer networks use object-oriented software technologies for distributed systems. They are highly scalable and modular, which enables efficient methods for software provisioning using commercial software products from different vendors. This approach facilitates systems that could be faster deployed with a higher flexibility and, hence, more cost efficiency.

The main principles (object-oriented analysis and design, distribution, decoupling of software components and separation of concern) have already been discussed in the Telecommunications Information Networking Architecture (TINA) consortium [25] during the past years.

However, to provide customers the same grade of service they are used to from telecommunication (e.g., necessary for voice transmission), requirements like the support of QoS, reliability, mobility, and security must be fulfilled. These requirements are a challenge to today's existing data networks (Internet). Moreover, there will be an increasing number of radio technologies, e.g., GSM, ANSI 136, cdmaONE, W-CDMA, TD-CDMA, IEEE 802.11, HiperLAN, that need to be supported in the near future.

From today's perspective, it is foreseeable that the mobile communications networks will evolve toward a client–server architecture. The following major requirements have to be fulfilled by the system, including the network and the terminals:

- QoS with real-time requirements (e.g., video streaming, voice, ...);
- reliability;
- access security;
- support of a wide range of access technologies and easy deployment of new ones (e.g., for operator differentiation);
- scalability of the components;
- self-optimizing networks (to drive down operation, administration, maintenance costs);
- distributed systems; objects may be located on any server in the network or the terminal;
- flexible in introducing new services, e.g., *ad hoc* networking (object oriented);
- secure software up- and download and configuration (of terminals and radio access points);
- storage requirements for thin clients;
- suitable for unreliable links (radio links that fail sometimes).

All these requirements should be addressed in architectures and software concepts that can handle distributed systems, concurrency, and the persistent storage of data throughout the network. These three dimensions are independent of each other and can be represented by a three-dimensional (3-D) coordinate system [26].

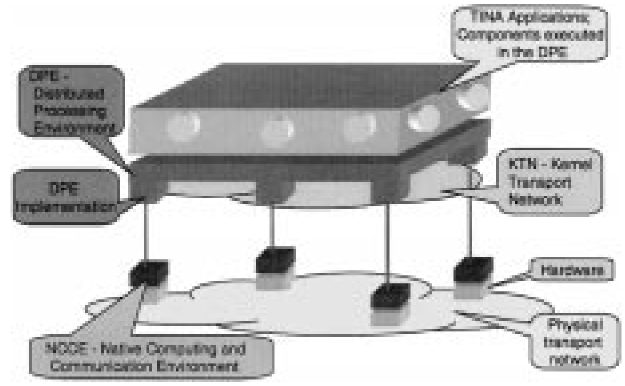

Fig. 11. Usage of the distributed processing environment (DPE), TINA-C.

Concurrency is the actual or apparent parallelism of processes in the same processing unit or in distributed processing units, and a major challenge is the synchronization of concurrent events in a distributed processing environment. JAVA threads is one possibility to implement parallel (concurrent) processes. Distribution means the logical or physical separation of objects. The simplest case for the communication among objects is the simple transfer of bits by sockets without its further interpretation. This is the baseline for advanced processing by adding semantics to the bit stream. More information about the socket technology in TCP/IP networks and its importance (e.g., for JAVA) can be found in [27] and [28]. Advanced concepts for distributed programming as object migration, distributed garbage collection, integration of CORBA and DCOM, mobile agents, and other features can be found in Voyager [29], which is also a middleware. Persistence is the reliable storing of data. This is important to cases where the mobile terminal is temporarily disconnected from the network during the data download. JAVA database connectivity (JDBC) enables the access to relational databases.

Listing these requirements leads to the point that a mobile middleware concept is needed that supports objects or agents to run and interwork in distributed systems. In this context, the support of mobility, QoS, security, support of different radio access technologies, and the usage of radio resources and their management can be interpreted as different objects in the overall network concept. Depending on the user's needs and the operator's requirements, several of these "objects" will be used. Objects that are located on the terminal device will interwork with the specific objects in the network. The basic principle of such an approach was discussed in the TINA architecture depicted in Fig. 11. Today, there are three main middleware technologies that support objects in distributed systems:

- COM/DCOM;
- JAVA RMI;
- CORBA that has been specified by the OMG (Object Management Group).

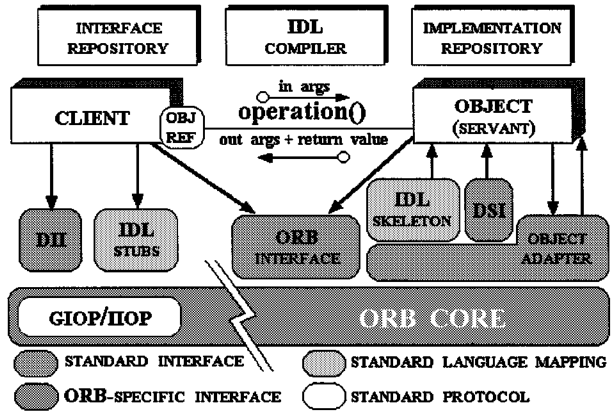

Fig. 12. CORBA ORB architecture.

The CORBA middleware is characterized by the following features.

- *Object oriented:* The definition of "object" within the object management architecture (OMA) is enlarged compared to the standard meaning of objects used in programming languages. The OMA uses "object" to an arbitrary uniquely identified entity within the architecture (see Fig. 12). Each object has a unique reference, even if it moves its location.
- *Object access transparency:* CORBA uses the same access mechanism to local and remote objects as well. The location of the server object is usually not visible to the client object.
- *Efficiency:* The object request broker (ORB) architecture allows efficient implementations, so that in case of local client–server object communication, the effort is only marginally increased compared to functional interfaces.
- *Open interfaces:* The ORB enables the interworking of programs of different vendors. This approach enables the user to use components from different vendors, tailored according to his individual needs. Software developers have the option to participate in tomorrow's software market and to complement products of large companies with specialized products.
- *Hardware, operating system (OS), and programming language independent:* Components of a CORBA program can be realized using different operating systems, hardware architectures, and programming languages.
- *Security concepts:* For reconfiguration, a reliable software download to the terminal or within the network must be accomplished. CORBA offers many different facets to facilitate this demand. For instance, mobile agents do not currently offer security concepts to fulfil these demands.[1]

Most of the characteristics except for the last one, which is CORBA specific, are also at least partly fulfilled by technologies like COM/DCOM and JAVA RMI. However, in contrast to JAVA RMI, COM/DCOM, or Voyager, which are developed and/or owned by one company, CORBA is an open standard that is supported from all important computer and software manufacturers.[2]

Due to intensive research work to develop real-time capable and QoS enabling CORBA during the past years, the required technologies, e.g., TAO ORB from the University of Washington in St. Louis [31]–[33], are now available for wireline networks.

Today, CORBA provides means for the support of the QoS for real-time applications [30]. Moreover, concepts to adapt applications according to the available and changing resources have been developed in the ACTS project OnTheMove [34], [35], e.g., in case of vertical handover. Further information can be also found in [36].

Today, CORBA is used in wireline data networks. What are the requirements that must be fulfilled by "wireless CORBA" to be suitable as a mobile middleware from the terminal to the network components and server in wireless mobile computing and communication? Addressing use cases, the OMG Telecom Domain Task Force released a whitepaper [8] that addresses the identified questions. A group within the OMG Telecommunication has started work on this topic with the target to finalize a first, out of a bundle, request for proposal (RFP) in May 2000. Other activities concerning this topic have been started in the 5th Framework Program of the European Commission for research projects [37] and are under consideration in the Software Defined Radio Forum (SDR-F) [38], [39].

For user applications, an extension of the current CORBA approach can be seen in the JINI concept, where *ad hoc* networking is facilitated, i.e., the user will be autonomously made aware of new services or network components in the user's environment. But JINI does not address the security aspects needed for the reconfiguration of terminals or network entities. More detailed information can be found in [40].

A general problem for security aspects is the IPv4 standard, which is used in the present TCP/IP world. There are some specific problems of wireless CORBA as, for example, the overload behavior of TCP/IP and the high load of the TCP headers. However, work has started in IETF to fix this problem.

There will be objects with different tasks in the future networks using the same CORBA technology. They will, however, have different requirements for storage, real-time capability, security, etc.

O&M objects (for network management) will have no specific requirements for storage and usually no real-time requirements, but high security and reliability requirements against fraudulent usage.

1) Minimum CORBA and MiniORBs: Very small and embedded devices are typically constrained with respect to resource usage such as main memory and CPU utilization. Full-blown CORBA implementations [41], however, offer a wide range of functionalities that are of limited use in such

[1]Details explaining Fig. 12 can be found at http://www.cs.wustl.edu/~schmidt/corba-overview.html.

[2]Approximately 800 companies out of a great variety of vertical markets, e.g., telecommunications, transport, aviation, media, are working within the Object Management Group [30].

devices. As a consequence, the Minimum CORBA specification, which essentially constitutes a part of the CORBA 3 standard, offers a reduced profile that even systems with low resources can support without losing interoperability to conventional CORBA systems. For this purpose, Minimum CORBA supports the full set of interface definition language (IDL) features. However, all dynamic aspects of CORBA have been eliminated. Examples for such aspects include the dynamic invocation interface (DII), the dynamic skeleton interface (DSI), the generic data type dynamic any and almost all parts of the interface repository. At the server side the number of policies as well as some of the options in the portable object adapter were removed.

Due to the assumption of a mobile middleware targeting all components from the terminal to the access system (e.g., node B and RNC) and further on to the server, it is needed to consider the required storage of CORBA and ORBs compared to the available storage on mobile terminals. Today's ORBs still require a huge amount of storage compared to the available one in today's mobile terminals.

In the near future, the wireless terminals will differentiate. Many new wireless devices will come to market, from the low-end wireless mobile phone (thin-client terminal) to handheld devices, from notepads up to notebooks. On the high end, there will be no real storage shortage due to fast increasing storage capacity. This will even meet the handheld devices up to a certain point. However, the mass market with the low-end devices and low-storage capacity will be for the near time a problem that must be solved. A way to fix this problem will be the location of a graphical user interface (GUI) on top of the terminal that interworks with an ORB located in the network.

However, for the support of a secure software download to future software radio terminals, it is advantageous to locate the ORB directly in the terminal. Investigations are needed to identify whether the required storage and the available functionality of so-called miniORBs could be provided and will be sufficient with the introduction of SDR terminals, or if some adaptations might be necessary or the definition of a microORB required.

Due to the differentiation of terminal capabilities and customer needs, it is required to develop and implement ORBs with different capabilities, from the full-fledged ORB, to the miniORB as it is implemented in the Netscape navigator, to a microORB (not yet defined), down to only a GUI (for low-end terminals) that interacts with the personal ORB located in the network.

2) Real-Time CORBA: Applications running on real-time systems must behave in a deterministic way with respect to their functional behavior as well as their QoS characteristics [42]. Hence, Real-time CORBA introduces many extensions and concepts to guarantee end-to-end predictability and to provide efficient resource management. Although the standard is based on the portable operating system interface (POSIX) real-time extensions, it does not prescribe the availability of a POSIX implementation.

An activity denotes a closely related sequence of actions that is triggered by an external event. Within an activity, different CORBA requests are dispatched or sent. In this context, the specification differentiates between three different states of an activity.

- In-transit: represents the transmission of a request.
- Static: represents a request that is cached in a queue.
- Active: denotes request dispatching by a thread.

A scheduling service is in charge of scheduling activities, i.e., scheduling the threads an activity consists of. Real-time CORBA only supports static scheduling strategies such as rate monotonic scheduling. Dynamic scheduling is (currently) not supported.

For initial configuration of real-time systems developers leverage the interface RT_CORBA::RTORB. Using additional APIs, they can create thread pools and change thread priorities. In order to guarantee platform neutrality, Real-time CORBA specifies a priority scheme that can be mapped to the native priority scheme of the underlying operating system. For converting priorities between these schemes, the interface priority mapping is available. Priority inversion is addressed using different strategies. For example, the RT-ORB transmits priority information from the client to the server using Internet inter-ORB protocol (IIOP) packets. This information enables the server to handle the incoming request with the appropriate priority. Special multiplexers are specified to synchronize resource access between parties. In addition, clients might establish multiple communication channels with the servers, where each of them gets assigned its own priority. They might also use dedicated connections that are not multiplexed, as well as specify timeouts for their requests.

B. Reconfiguration of Software Radio Terminals with Wireless CORBA

The client–server methodology for distributed objects is adopted, which can be well described by the CORBA approach.

CORBA provides an open, broadly applicable, and platform-neutral technology that uses object-oriented techniques to hide most of the accidental complexity introduced by distribution.

With the emergence of wireless network technologies, such as GSM, IMT2000/UMTS, and WLAN, it is important that CORBA can be extended to accommodate networks that contain wireless links. Mobile terminals generally have access to fewer local resources than is possible with stationary devices, both in terms of processing power and information. Users increasingly expect to be able to access sophisticated and powerful services from their mobile devices that they currently access from stationary devices. As a consequence, the ability to use services of remote objects become more and more urgent.

The need to change the air interface or basic functionalities of terminals is driven by the assumption that some services might be temporarily or permanently available with a certain air interface.

Wireless CORBA is currently under standardization to consider the wireless scenarios, i.e., temporary disconnections and terminal mobility. It must extend the current CORBA by taking into account the aspects shown in Fig. 13. Wireless

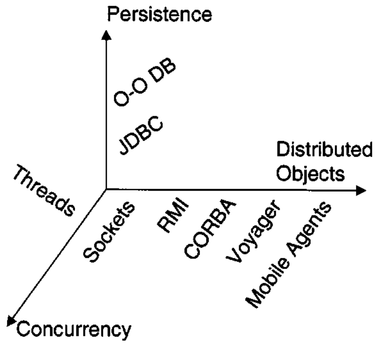

Fig. 13. Architectural and software aspects for future network concepts.

CORBA must be able to cope with migration of objects from one platform to a remote platform, asynchronous RMIs, etc.

Wireless CORBA should be the concept to integrate the handling with distributed objects and concurrent processes, as depicted in Fig. 14.

The virtual processor is an active object and in parallel the run unit for several objects, i.e., the virtual processor controls the objects allocated to it. The middleware can be seen as a virtual blackplane and establishes the communication between remote objects in the terminal, node B, or RNC and supports migration of objects, i.e., objects can roam between different network entities. Depending on the available resources in a terminal, applications can fully run on the terminal or in the network, and remote accesses between the terminal and network must be established.

Dejay [9] is a dialect of JAVA that simplifies the development of distributed software applications. Concurrent processes, distribution of objects and persistence aspects need to be addressed and require the learning of many different techniques like JAVA Threads, JAVA RMI, CORBA, and OODB. Wireless CORBA needs to address all these issues but does not have to lose its security aspects needed for a reliable reconfiguration.

The link between terminal and node B is established via the mobile wireless inter-ORB protocol (MWIOP) and can be seen as extension of the generic interORB protocol (GIOP). The current GIOP as it exists today must be enhanced accordingly, and the new MWIOP shall take into account the wireless requirements to run CORBA. CORBA offers many security facets to make reconfiguration of terminals or networks reliable, and these facets need to be translated into the wireless world. Additionally, CORBA facilitates the interworking of proprietary equipment distributed across the network or within one cabinet.

The software download can be seen as a special application that needs a very high reliability for the reconfiguration process, which can affect the whole functionality within a terminal or network element. In the European Research project TRUST [43], the reliable download is one of the important research areas for terminals, but similar security demands are needed for other network elements such as, e.g., base stations.

To reconfigure basic functionalities within a terminal or base station, a functional framework for a software radio terminal can be outlined, as in Fig. 15.

Proprietary objects as, for instance, the physical layer, communicate via a standardized interface description language (IDL) over a software bus, which consists of the ORB. The communication is related to the reconfiguration management of these objects, but in the future it might be possible to transfer the data stream via the ORB too, if the performance demands are solved. Special objects are needed to handle the download, the installation, and the security check of new software.

Fig. 16 shows the interaction between an application server or a server, provisioning software for a new air interface (later on called "wave server") and a mobile terminal. Here, a JAVA VM (virtual machine) is assumed for the mobile station execution environment (MExE). The Internet paradigm involves pulling JAVA byte code from a web server to a front-end deployment platform on a terminal.

Combining JAVA and CORBA has several important benefits.

- The CORBA three-tier model combined with the platform independence of JAVA supports the pull paradigm for application distribution, greatly simplifying maintenance of the application-installed base. Furthermore, an application architecture based on CORBA is far more scalable than the traditional two-tier client–server model. Middle-tier logic can be partitioned and dynamically distributed, thereby moving processing closer to data sources and supporting load balancing for downloads.
- Since CORBA facilitates language independence and location transparency, JAVA objects can communicate with other distributed objects. This provides an effective means of web-enabling applications without incurring the bottlenecks experienced with HTTP- and CGI-based approaches.

The client objects for the radio network layer (RNL) on the terminal consist of the physical layer, RLC/MAC, and the RRC layer. The air interface software server, i.e., wave server in the network carries all information for the terminal reconfiguration, which can be established by the software downloader object located in the terminal. The air interface software is hosted on a wave server, which is contacted on demand.

The following phases of a reconfiguration for a terminal can be abstracted into the following phases or states:

- initial software download onto a terminal, e.g., by the SIM Application Toolkit (SAT);
- discovery of new mobile air interface servers, service or applications servers;
- binding to remote servers for initializing the software download;
- follow-up software download onto a terminal;
- installation and activation of new air interface, services, or applications;
- deinstallation of previously accepted air interface, services, or applications;
- remote method invocation;
- unexpected disconnections (and reconnections).

1) Installation and Activation of a New Air Interface: Installation can be static or dynamic. Static installa-

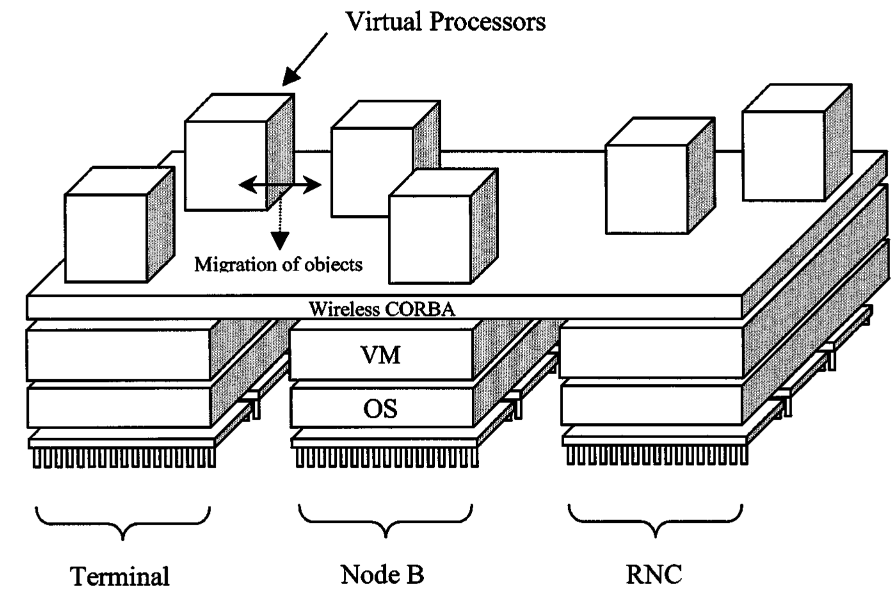

Fig. 14. Abstraction of a future middleware for wireless systems with JAVA.

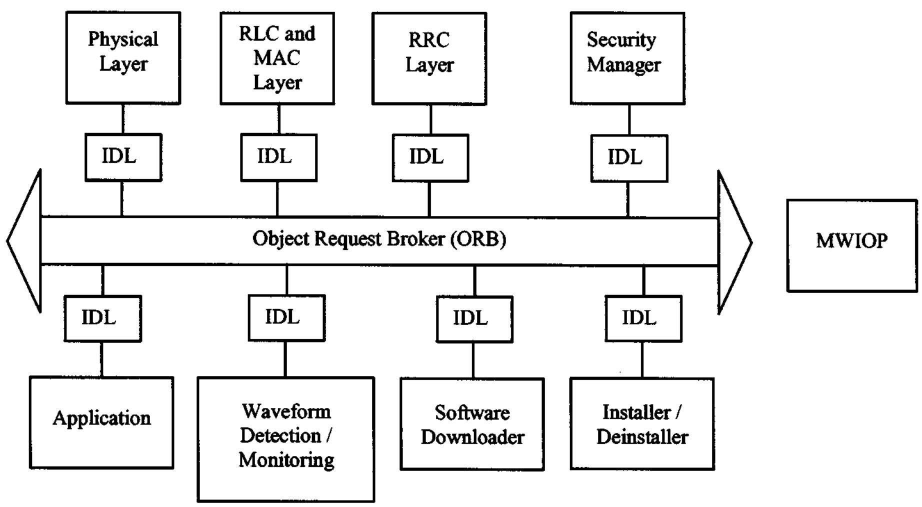

Fig. 15. Functional framework for software radio terminals applying wireless CORBA.

tion and the activation of statically installed applications means, e.g., to do it via the subscriber identity module (SIM card). Further, one has to distinguish whether the download and reconfiguration process are carried out in idle or active terminal mode. In any case, client objects need to get registered first to the server in the network, before the download can start.

Software to be dynamically loaded will be of great importance for terminals for the following reasons.

- Resources available at mobile terminals will, also in the future, be relatively limited.
- Mobile terminals can roam from network to network where different applications are required or offered to the user.
- The faster and faster churn rate of new application software in general and the need to try to take the burden of installing and maintaining new air interface features or user applications, e.g., optimized channel coding schemes for applications away from the users of mobile terminals.

With the object-by-value (OBV) specification, loading software dynamically should be practical. This solves the problem of getting software dynamically over the over-the-air reconfiguration (OTAR). However, this does not fix the activation problem after a security check. When software for a new air interface is downloaded, it must be configured and tested before it can be used. This involves looking up information on capabilities (e.g., interference situation on measured carriers), terminal resources, and user preferences of the local environment before a download takes place. It might also require fetching software for the new air interface if they are not available locally. For example, if a user in Europe intends to travel to the United States, he might demand to download a new air interface standard before entering the aircraft. What is needed is a standard runtime environment that prompts dynamically loaded CORBA soft-

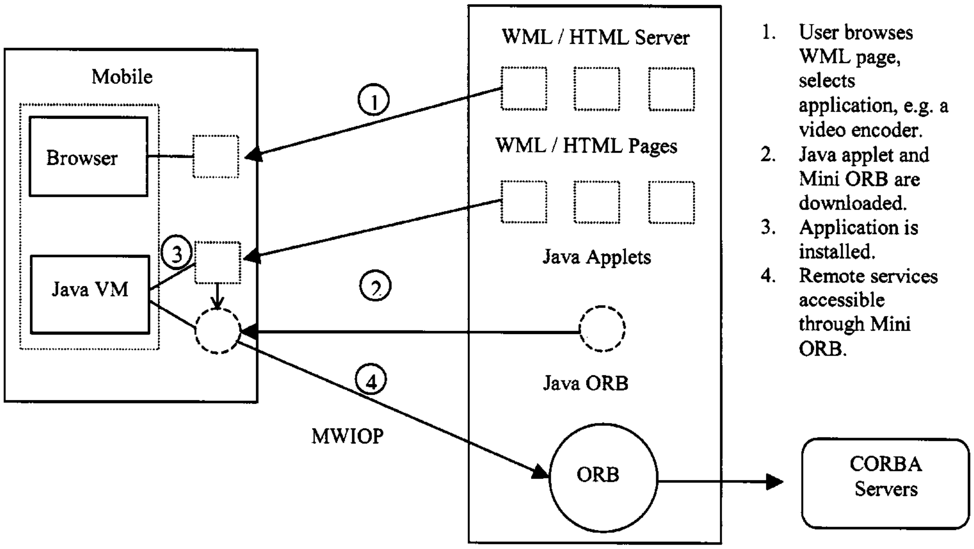

Fig. 16. JAVA ORB on a software radio terminal.

ware to prepare itself after a successful security check and then activates (reconfigure) the terminal, when appropriate.

There are two main aspects to software download: the transport mechanism and the process. The transport mechanism is how the software is entered into the terminal. The process covers for the actions and activities that must take place for the software to be successfully installed into the terminal. Dynamic reconfiguration over the air offers in principle two alternatives: interactive download (dedicated data channel) and broadcast download (common data channel).

Before a download can take place, an alternative air interface must be identified by the terminal. In principle, blind versus bulletin system identification can be addressed. Blind identification means that no prior knowledge of available air interface is given and bulletin system identification utilizes systemwide signaling channels, which gives a pointer to where a local download channel can be located. Download channel schemes need to be specified, i.e., where software is to be transferred to the terminal.

The pure JAVA world has the JAVA Embedded Server API to do this. It also provides some other useful features such as allowing software to be automatically and transparently upgraded when a new mode feature or software application becomes available. The CORBA solution might leverage the work on CORBA mobile agents if this can be made lightweight enough to be used on mobile terminals.

2) Discovery of New Air Interface Servers: The identification and monitoring of new air interfaces is carried out by the mode (air interface) identification and monitoring object in the terminal. A successive software download via the air requires to determine the wave server by getting an initial reference to the naming or trader service from the miniORB at the terminal. The naming or trader service are CORBA services that have a reference to the actual server address. Wireless CORBA must be extended in such a way that a simple, standard scheme for advertising and finding air interface or application software servers enables a fluid change of mobile network cells or whole mobile networks. Terminals that carry, e.g., an out-of-date manufacturer software shall be easily updated when required and appropriate.

In particular, the access gateway should provide a naming and/or trading service, the initial reference of which the miniORB on the terminal gets during the initial access procedure.

3) Binding to Remote Servers for New Air Interface: There are binding issues. First, there is the problem of binding to a wave or software server that has a network address that may change during the download. Therefore, some sort of relocatable object reference may be required for this. Another binding issue involves the choice of network download channels when a mobile terminal has more than one choice for the channel that can be used between a client and air interface software server for download.

Presumably, a server would advertise the available means of interacting with it in the tagged profile data in its MWIOP. After this, the client would make the decision about which channel to use and then use the service context list to signal any specific references back to the network. It may be possible that the channel may change due to handover reasons or an interruption time may delay the download.

Yet another binding issue is the versioning, i.e., different configuration software version for possibly different mobile terminals. In traditional CORBA systems on fixed networks, that can be centrally administered. For simplicity reasons, this method should be used also in wireless networks.

4) Remote Method Invocation: The software download will be invoked by the terminal, if the radio access network permits it. The download can be considered as a special user application that may tolerate higher data delay but requires a very low bit error rate (BER). In some cases, it will be possible to allocate network or server resources in such a way that the configuration software is faster or more reliable downloaded. To take advantage of this, CORBA applications require a mechanism for expressing their QoS requirements, not only for the download application, but also for others.

In most cases, it would probably be appropriate to configure these requirements for a particular user or terminal using network QoS policies that are attached to the miniORB through an "ORB QoS Service." If the QoS is insufficient, CORBA needs to give the download mechanism access to the QoS information about the current characteristics of the computing and communication environment and an easy way to modify the download channel or to change the network. In RT-CORBA, end-to-end QoS enabling mechanisms are already specified.

5) Unexpected Disconnections and Recovery: Unexpected disconnections are inevitable with wireless radio systems. If an interruption occurs during a download, the client–server needs a global knowledge and control of the download state in order to determine what resources are still in use and to free them appropriately in the network and on the terminal, if the download cannot be resumed within a certain time limit. A garbage collector is then responsible to free unnecessarily occupied resources.

6) Location-Dependent Software Download: The location service supports CORBA mobility-aware software download. The download can ask for the current location for a client. The mobility server updates the data in the location register when a terminal has moved to a new location for introducing location-dependent services or applications.

7) Authentication Service for Software Download: The authentication service supports the download mechanisms that require a terminal to authenticate itself before allowing it to download certain software. The downloader asks the authentication server to check whether an object is allowed to access a software to be downloaded to the terminal. The authentication server interacts with the terminal (e.g., exchanging authentication keys) to secure that the user (terminal) is the one it claims to be.

V. Conclusion

Mobile multimedia applications have already been introduced into evolved second-generation mobile radio systems. The improved wide-band and more flexible radio interfaces of third-generation mobile radio systems provide even more opportunities for mobile multimedia services and applications. Wireless communication systems beyond the third generation will combine many optimized access systems for special purposes that are integrated on a common flexible network platform. These different access systems complement each other in an efficient and optimized way from the user's perspective. Depending on the selected services, the available access systems, and the bearer capabilities, this new system will select the most appropriate access. The different access systems are allocated to different cell layers with respect to cell size, coverage, and mobility for globally optimized seamless service provision. The flexible platform of the medium-access system and the core network will be based on IP technology with transparent transmission to ensure flexibility for all involved players in the new deregulated and liberalized communication environment. In addition, the concept will use algorithms for reconfigurability, self-optimizing networks, and automatic planning so that new network entities and access systems can easily be added.

Key issues of the new concept are the interworking of a variety of access systems on the common platform by horizontal and vertical handover, service negotiation for seamless service provisioning, and global roaming. A new medium-access system and the IP-based core network will handle the interworking and mobility management.

There are many technical challenges that have to be solved to realize this vision. This requires extensive international research activities in the areas of access system improvement, optimization of IP for radio transmission and mobility management as well as the implementation of multimode, multiband terminals, and, finally, software radio terminals and network entities. International standardization will play an important role to implement this concept on a global basis to ensure global roaming and seamless service provisioning. Besides the conceptual approaches for capacity increase and improved QoS, the reconfigurability of terminals and network components is seen as a major technical challenge. Adequate concepts have been presented, which can be the next major leap forward in the integration of heterogeneous networks for traditional mobile networks as well as computer networks. The demands for new concepts for distributed processing in a wireless environment have been outlined. These new demands can be translated to the integration of managing distributed objects, concurrency, and persistence in a wireless world. In particular, the reliable software download must be realized by, e.g., a new middleware, which takes into account the already available security concepts in CORBA. The requirements on wireless CORBA for a software download have been described. The reliable download is inevitable for the radio network layer (physical layer, MAC/RLC, and RRC), but for crucial user applications, such as home banking, the same security demands must be raised. Capacity enhancing features like multiuser detection and smart antennas will also play an important role.

Finally, economic aspects will determine which access system elements will be implemented. However, the concept should be open and flexible so that new access systems can be introduced later to improve the service to users continuously.

Acknowledgment

The authors would like to thank M. Stal (Siemens ZT), H.-G. Baumgarten (Siemens ZT), M. Frank (Siemens ICN M), and J. Hauck (Siemens ICN M) for their useful discussions.

References

[1] F. Ménard, D. Isenberg, and T. M. Denton, Consultants. Netheads versus Bellheads. Final Rep., Ottawa, Canada. [Online]. Available: http://www.tmdenton.com/netheads3.htm

[2] K.-U. Stein, Zur technoökonomischen Evolution der Telekommunikation, Nachrichtentechnische Zeitschrift, Organ der Information-stechnischen Gesellschaft im VDE(ITG), ntz 9/99, pp. 72–75.

[3] WAP Forum. [Online]. Available: http://www.wap.org

[4] UMTS Forum. [Online]. Available: http://www.umts-forum.org/

[5] W. Mohr and W. Konhäuser, "Access network evolution beyond third generation mobile communications," *IEEE Commun. Mag.*, 2000.

[6] DVB-Forum. [Online]. Available: http://www.dvb.org/
[7] M. Haardt and W. Mohr, "The complete solution for third generation mobile communications: Two modes on air—One winning strategy," presented at the Proc. IEEE Int. Conf. Third Generation Wireless Communications, San Francisco, CA, June 2000.
[8] "White paper on wireless access and terminal mobility in CORBA," OMG Telecom Domain Task Force, OMG Document: Telecom/98-11–09.
[9] Distributed JAVA: Dejay. [Online]. Available: http://www.dejay.org
[10] G. Giannakis, P. A. Anghel, Z. Wang, and A. Scaglione, "Generalized multi-carrier CDMA for MUI/ISI-resilient uplink transmission irrespective of frequency-selective multipath," in *Multi-Carrier Spread Spectrum & Related Topics*, K. Fazel and S. Kaiser, Eds. Norwell, MA: Kluwer, 2000.
[11] C. Zhao, T. Katura, M. Kisimoto, T. Ueda, T. Matukawa, and H. Okada, "Network connectivity of layered selforganizing wireless networks," presented at the Proc. 6th Int. Conf. Computer Communications Networks, Las Vegas, NV, Sept. 1997.
[12] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, 1997.
[13] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, 1998.
[14] M. Tangemann, C. Hoek, and R. Rheinschmitt, "Introducing adaptive array antenna concepts in mobile communication systems," in *Proc. RACE Mobile Telecommunications Workshop*, vol. 2, Amsterdam, The Netherlands, May 1994, pp. 714–724.
[15] M. Vollmer, J. Goetze, and M. Haardt, "Efficient joint detection techniques in the frequency domain," presented at the Proc. IEEE Int. Conf. Third Generation Wireless Communications, San Francisco, CA, June 2000.
[16] P. W. Baier, A. Papathanassiou, and M. Weckerle, "Recent results on the benefits of adaptive antennas in TD-CDMA mobile radio systems," in *Proc. IEEE/IEE Int. Conf. Telecommunications (ICT '99)*, vol. 1, Cheju, South Korea, June 1999, pp. 399–404.
[17] C. Brunner, M. Haardt, and J. A. Nossek, "On space-time rake receiver structures for WCDMA," presented at the Proc. 33rd Asilomar Conf. Signals, Systems, Computers, Pacific Grove, CA, Oct. 1999.
[18] M. Haardt, C. F. Mecklenbräuker, and M. Vollmer, "Adaptive antennas for third generation mobile radio systems," presented at the Proc. MICRO.tec 2000, Hannover, Germany, Sept. 2000.
[19] T. Asté, P. Forster, and L. Féty, "Downlink beamforming avoiding DOA estimation for cellular mobile communications," presented at the Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), Seattle, WA, 1998.
[20] H. Bölcskei, R. W. Heath, and A. Paulraj, "Blind channel identification and in OFDM-based multi-antenna systems," *IEEE Trans. Signal Processing*, submitted for publication.
[21] H. Bölcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based multi-antenna systems," presented at the Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 2000), Istanbul, Turkey, June 2000.
[22] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.
[23] H. Sampath and A. Paulraj, "Transmit and receive optimization for high data rate wireless communication using multiple antennas," presented at the Proc. 33rd Asilomar Conf. Signals, Systems, Computers, Pacific Grove, CA, Oct. 1999.
[24] *IEEE Commun. Mag.*, vol. 37, no. 2, pp. 82–112, 1999. (Several articles on software defined radio).
[25] TINA (Telecommunications Information Networking Architecture) Consortium. [Online]. Available: http://www.tinac.com
[26] M. Boger, *JAVA in Distributed Systems*, 1999.
[27] M. Hughes, C. Hughes, M. Shoffner, and M. Winslow, *Communicating Sequential Processes*. Greenwich, CT: Manning, 1997.
[28] P. Neimeyer, *Core JAVA Networking*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
[29] Voyager from Objectspace. [Online]. Available: http://www.objectspace.com
[30] Object Management Group, OMG. [Online]. Available: http://www.omg.org
[31] D. C. Schmidt, D. L. Levine, and C. Cleeland, "Architectures and patterns for high-performance, real-time ORB endsystems," in *Advances in Computers*, M. Zelkowitz, Ed. New York: Academic, 1999.
[32] S. Mungee, N. Surendran, and D. C. Schmidt, "The design and performance of a CORBA audio/video streaming service," presented at the Proc. 32nd Hawaii Int. Conf. System Systems (HICSS), Jan. 1999.
[33] C. O'Ryan, F. Kuhns, D. C. Schmidt, O. Othman, and J. Parsons, "The design and performance of a pluggable protocols framework for real-time distributed object computing middleware," presented at the Proc. IFIP/ACM Middleware 2000 Conf., Pallisades, NY, Apr. 3–7, 2000.
[34] ACTS Project OnTheMove. [Online]. Available: http://www.sics.se/~onthemove/
[35] T. Wierlemann, T. Kassing, K. David, and J. Hjelm, "OnTheMove field trials: Evaluating mobile middleware for multimedia applications," presented at the ACTS Mobile Communications Summit, Aalborg, Denmark, Oct. 7–10, 1997.
[36] B. Keller, A. Sang-Bum Park, J. Meggers, G. Forsgren, E. Kovacs, and M. Rosinus, "UMTS: A middleware architecture and mobile-API approach," *IEEE Pers. Commun.*, vol. 5, Apr. 1998.
[37] Information to the Fifth Framework on the Information Society Technologies Programme of the European Commission. [Online]. Available: http://www.cordis.lu/ist/overview.htm
[38] B. Stephen and D. Murotake, "Software-defined radio moves into base-station designs," *Wireless Syst. Des.*, Nov. 1999.
[39] SDR Forum. [Online]. Available: http://www.sdrforum.org
[40] JINI Information. [Online]. Available: http://www.sun.com/jini
[41] M. Stal, *CORBA3.0—Part 1: Komponenten—Episode 3: ix-Magazin*. Germany: Heise-Verlag, Mar. 2000.
[42] ——, *CORBA3.0—Part 2: Internet and Quality of Service—World Wide CORBA: ix-Magazin*: Heise-Verlag, Apr. 2000.
[43] Transparent Reconfigurable Ubiquitous Terminal (TRUST), EU Research Project. [Online]. Available: http://www.ist-trust.org

Reinhard Becher received the M.S. and Ph.D. degrees in theoretical physics from the University of Regensburg, Regensburg, Germany, in 1984 and 1989, respectively.

In 1989, he joined the Mobile Network Division at Siemens, Munich, Germany, where he developed call processing software for the GSM switching subsystem. In 1993, he joined the System Engineering Division, where he was responsible for the conceptual work on intelligent network aspects in GSM mobile networks. His main tasks were concepts for fixed-mobile (GSM) integration, the standardization of CAMEL (intelligent network architecture in GSM networks), as well as the Virtual Home Environment concept. Since 1999, he has been serving as Director for Standardization and Network Technologies in the Chief Technical Office of Siemens Communication on Air.

Dr. Becher received the Siemens public networks division Inventor of the Year Award in 1998.

Markus Dillinger (Member, IEEE) received the Dipl. degree in telecommunications from the University of Kaiserslautern, Kaiserslautern, Germany, in 1990.

In 1991, he joined the Mobile Network Division at Siemens, Munich, Germany, where he developed call processing software for the GSM base stations. Later, he joined the System Engineering Division, where he was responsible for the evaluation software for the Siemens Channel Sounder measurement equipment. In addition, he was responsible for the definition of the technical requirements and specification on the GSM base station line interfaces. From 1995 on, he worked on the definition of the third mobile radio generation in the European Research Project FRAMES, and in 1999, he was appointed Technical Manager of FRAMES. Since January 2000, he has been the Project Leader of the European Research Project TRUST for Software Radio. He holds many patents and has published many articles in the field of WCDMA and software radio.

Martin Haardt (Senior Member, IEEE) received the M.S. degree from the Ruhr-University Bochum, Bochum, Germany, in 1991 and the Ph.D. degree from the Technical University of Munich, Munich, Germany, in 1996, both in electrical engineering.

From 1989 to 1990, he was a Visiting Scholar at Purdue University, West Lafayette, IN, sponsored by a fellowship of the German Academic Exchange Service (DAAD). From 1991 to 1993, he was with Siemens AG, Corporate Research and Development, Munich, conducting research in the areas of image processing and biomedical signal processing. From 1993 to 1996, he was a Research Associate with the Institute of Network Theory and Circuit Design at the Technical University of Munich, where his research was focused on smart antennas. In 1997, he joined Siemens Mobile Networks, Munich, where he was responsible for strategic research for third-generation mobile radio systems. Since 1998, he has been serving as Director for International Projects and University Cooperations in the Chief Technical Office of Siemens Communication on Air. He also teaches at the Technical University of Munich. His current research interests include wireless communications, array signal processing, digital signal processing, and numerical linear algebra. He is a contributing author to *The Digital Signal Processing Handbook* (CRC Press and IEEE Press, 1998).

In 1997, Dr. Haardt received the Rohde & Schwarz Outstanding Dissertation Award for his Ph.D. dissertation. In 1998, he was the corecipient of the Mannesmann Mobilfunk Innovations-Award for outstanding research in mobile communications and received the ITG best paper award from the Association of Electrical Engineering, Electronics, and Information Technology (VDE).

Werner Mohr (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of Hannover, Hannover, Germany, in 1981 and 1987, respectively.

From 1987 to 1990, he was a Senior Engineer at the Institute of High-Frequency Technology, University of Hannover. Since 1991, he has been with Siemens AG, Mobile Network Division, Munich, Germany. He has been responsible for propagation measurements and channel modeling, and he has been involved in the European RACE-II Project ATDMA. From 1995 to 1996, he was active in ETSI SMG5 for standardization of UMTS. Since December 1996, he was Project Manager of the ACTS FRAMES Project until the project finished in August 1999. He was Director of Strategic Pre-Development up to September 1998, and since October 1, 1998, he has been Vice President Pre-Engineering in the Chief Technical Office of the Communication on Air business area of Siemens ICM. Currently, he is involved in the fifth framework program of the European Commission.

In 1990, Dr. Mohr received the ITG Award from the Association of Electrical Engineering, Electronics, and Information Technology (VDE). He is member of VDE.