Moving towards Seamless Mobility—State of the Art and Emerging Aspects in Standardization Bodies

Marc Emmelmann, Sven Wiethoelter, Andreas Koepsel, Cornelia Kappler, and Adam Wolisz

Abstract— The challenge to provide seamless mobility in the near future emerges as a key topic in various standardization bodies. This includes first of all the support of seamless handover between homogeneous networks. Distinct technologies — such as IEEE 802.11WLANs (Wi-Fi) and IEEE 802.16 MANs (WiMAX)—have recently augmented such support to existing standards to enable seamless homogeneous handover. Cellular networks, in contrast, already included this inherently from the start. Currently considerable effort goes into coupling of different radio access technologies. Therefore, the second key topic in standardization is seamless heterogeneous handovers. IEEE, IETF, as well as 3GPP consider different approaches towards architectures and protocols enabling seamless mobility management. In this work, we discuss recent and on-going standardization activities within IEEE, IETF, and 3GPP towards seamless homogeneous as well as heterogeneous mobility support.

Index Terms— Seamless mobility, Standardization, IEEE, Wi-Fi, 802.11, 802.11r, 802.11k, 802.11F, IAPP, WiMAX, 802.16, 802.16e, 802.21, Media Independent Handover (MIH), IETF, Host Identity Protocol (HIP), Stream Control Transmission Protocol (SCTP), SIGTRAN, Dynamic Address Reconfiguration (DAR), MIPv6, MIPSHOP, Fast MIPv6 (FMIPv6), Hierarchical MIPv6 (HMIPv6), SeaMoby, Context Transfer Protocol (CXTP), Candidate Access Router Discovery (CARD), DNA, NETLMM, MOBIKE, HOKey, MONAMI, 3GPP, System Architecture Evolution (SAE), Unlicensed Mobile Access (UMA)

I. Introduction

Wireless access technologies as well as the number of mobile devices have continuously grown over the last decades. The importance of mobility support continuously shifts away from mere nomadic networking towards mobile networking. The latter enables users to maintain their application session while moving within a single or among several access technologies.

Even though mobile networking is possible today, it cannot provide service continuity completely. For such seamless mobility an ongoing application session has to be maintained continuously such that an acceptable quality of service (QoS) perceived by a user is sustained.

As seamless mobile networking is further and further evaluated by its ability to support QoS-sensitive applications — i.e.

voice or video conferencing — while maintaining a secured connection, ongoing standardization efforts focus on three mobility aspects:

1) providing seamless handover for homogeneous technologies,
2) providing seamless handover among different access technologies, and
3) integrating different access networks and technologies under a common IP backbone.

Seamless homogeneous handover has been an integral aspect of cellular networks whereas IEEE 802-based wireless networks, i.e., IEEE 802.11(Wi-Fi) and IEEE 802.16 (WiMAX), are currently working on amendments providing such schemes. For heterogeneous handover, the IEEE 802.21 Media Independent Handover working group progresses to establish generic SAPs and service primitives which allow to trigger and indicate the need for handover. Meanwhile, 3GPP works in the 3G System Architecture Evolution context investigating schemes to establish a unified mobility concept. Finally, the IETF targets seamless mobility among heterogeneous access technologies connected via a common IP-based layer-3 infrastructure.

This article provides an overview of the most recent and ongoing standardization efforts enabling seamless mobility in both, homogeneous and heterogeneous environments. Section II starts with a classification of handover phases and sketches some of the involved QoS- and security-related challenges. Afterwards, solutions currently under discussion are outlined according to standardization bodies: Section III presents IEEE's work to enable seamless mobility in a homogeneous IEEE 802.11 (Wi-Fi) and IEEE 802.16 (WiMAX) environment as well the current status of IEEE 802.21, the Media Independent Handover group. Section IV summarizes ongoing work of the IETF regarding mobility support, i.e., components of SCTP and HIP, combinations of Hierarchical Mobile IPv6 and Fast Mobile IP, and presents approaches of the SEAMOBY, DNA, NETLMM, MOBIKE, HoKEY, and MONAMI6 working groups. Finally, Section V lays out mobility support in 3GPP and highlights goals for the System Architecture Evolution (SAE). Trends to couple 3GPP networks with IEEE-based wireless networks such as WLAN and WiMAX conclude the paper.

II. CHALLENGES FOR PROVIDING SEAMLESS MOBILITY

In general, a handover process can be subdivided into several phases, each of them requiring optimization in order to provide seamless mobility. The involved "base functions" can be coarsely denoted as discovery and detection of available network attachment points, handover decision and criteria, as well as link / connection re-establishment. [1], [2]

If not being handled in parallel to other actions, the *discovery and detection of network attachment point* is the most time consuming phase. Some technologies traditionally require scanning of several available frequencies in order to detect available radio access technologies (RATs) as well as, regarding higher layers, the reception of information on available access routers. The optimization of this phase should work towards reducing the time-span of this phase as well as assuring the credibility and integrity of detected NAPs, which in the case of finding a rogue NAP could lead to man-in-the-middle attacks. [3], [4]

Even though the *handover decision and involved criteria* are not in the scope of standardization bodies, they remain a challenging research issue, e.g., in a mobile environment in which the terminal's velocity influences both, handover failure probabilities as well as requirements towards the minimal overlap of adjacent radio cells for seamless handover. [4], [5]

The *link connection / re-establishment* phase is usually responsible for (re-)negotiation of resources. It involves a handshake at layer-2 and — in case of an IP subnet change — certain actions for address (re-)configuration either at the end host or at entities within the network. For such an address reconfiguration, the change of the new NAP has to be signaled to network components involved in the communication which eventually leads to a re-direction of the communication path. [2] As a change of the NAP usually requires to (re-)negotiate cipher keys for secure communication between nodes, this time consuming operation should also be optimized by, e.g., distributed key hierarchies. Involved security risks, i.e., man-in-the-middle or lack of privacy [3] are still challenging issues.

III. IEEE

A. IEEE 802.11

For IEEE 802.11 devices — employing a decentralized, CSMA/CA-based medium access — mobility is only supported in infrastructure mode in which several stations (STA) and an access point (AP) form a basic service set (BSS). In order to enlarge the wireless coverage area, a distribution system (DS) may connect several BSSs forming an extended service set (ESS). Moving from one AP's coverage into another's implies detecting the loss or degradation of the current connection, determining an AP to handover to, and establishing a new layer-2 connection with the new AP, i.e., authentication and association. As these steps may last several seconds [6] means for providing seamless mobility support were amended to the standard.

Algorithms on how to detect the loss or degradation of an ongoing connection while moving are not standardized but may be based on, e.g., three consecutively missing beacons, five consecutively failed transmissions [7] or SNR measurements retrieved from the physical layer (PHY). Even though the detection phase will remain proprietary, IEEE 802.11k amends radio resource measurement schemes facilitating decision algorithms by introducing a measurement pilot frame: a compact management frame periodically transmitted by an AP with a period much smaller than the beacon interval.

Compared to the beacon, the pilot provides a minimal set of information including its employed transmission power and noise floor at the AP. In combination with the SNR experienced at the receiver, it allows a link margin calculation suitable for transition decisions. Additionally, IEEE 802.11k allows to automatically trigger reports, e.g., if the received channel power falls below a certain threshold, as well as to exchange location configuration information both enabling link status- or position-based handover decisions. [8]

The most time-consuming phase during handover is the scanning phase [6] , which is significantly reduced by the above mentioned pilot frame and neighborhood information reports. The former's small transmission interval reduces the time spent by a STA on each channel during passive scanning. The latter contains information on validated neighbor APs that are members of ESS and allows scanning on selected frequencies only or even avoids scanning at all. It should be noted that the amendment does not specify means on how to generate that list but reveals one possible approach: a STA scans for APs, builds a local neighbor report, and exchanges it with the AP. [8]

The IEEE 802.11r fast BSS transition amendment optimizes the number of exchanges required to establish an authentication between the STA and new AP and suggests to employ IEEE 802.11k schemes to reduce scanning times. Instead of conducting an authentication "over the air" as in legacy IEEE 802.11, a remote request broker (RRB) is introduced at each AP. Instead of addressing the target AP, the STA directs its authentication request to the RRB which in turn encapsulates and forwards it to the target AP's RRB "via the DS". The latter interacts with the new AP's STA management entity to establish authentication. Besides, a STA may request resources at the new AP via the DS using the RRB. This allows the MT to uphold simultaneously an active communication channel via the old AP while checking and finally deciding for a new AP. In advance, IEEE 802.11r introduces optimized message exchanges establishing security by key forwarding and distribution. This is achieved by including an hierarchical key structure which is derived during the initial session set-up from an Extensible Authentication Protocol (EAP) master session key (MSK) [9]. The highest key is the Pairwise Master Key (PMK) which is cached at several APs and thus avoids the need for a new, full, and secure authentication upon each L2-handover. During a handover, the AP uses the PMK to derive a new, cryptographically separated key for — the Pairwise Transition Key (PTK) — for each session [10], [11].

After the handover, the old AP might still have packets addressed to the MT in its buffer. IEEE 802.11F [12] provided a recommended practice for an inter access point protocol which allowed the new AP to trigger the old AP forcing the

latter to forward these packets.[1] Additionally, IEEE 802.11r provides a de-authentication via the DS to release resources at the old AP. [10]

Up to now, IEEE 802.11 based handover schemes are entirely mobile controlled. The wireless network management working group IEEE 802.11v discusses a paradigm shift towards supporting a network directed handover allowing to achieve, e.g., load balancing between APs. [13], [14] Additionally, mechanisms to dynamically adjust individual handover policies at a STA, e.g., by the AP, are under discussion. [15], [16]

B. IEEE 802.16

IEEE 802.16 networks provide centralized broadband wireless access. The BS controls the (mobile) subscriber stations—(M)SS—employing a combination of time division multiple access (TDMA) and demand assigned multiple access (DAMA). The downstream can be based upon continuous time division multiplexing (TDM) or slotted, TDMA-like bursts. Similar to IEEE 802.11, the handover process can be divided into the detection of link's degradation or loss, the exploration of possible target BSs, as well as authentication and association. Additionally, due the strictly timed media access scheme, SSs have to synchronize themselves to the BS and have to adjust the employed transmission power (ranging process). IEEE 802.16e amends a mobility support already optimized in term of reduced handover delays.

To detect the need for handover, BSs may mandate SSs to continuously monitor the carrier-noise-interference-ratio (CNIR) and report its mean and its standard-deviation via a prioritized fast feedback channel. This information may serve as an input for handover algorithms which are not standardized.

To establish a knowledge on neighboring BSs, SSs may periodically scan for neighbor BSs. Therefore, the SS may request a time interval reserved for scanning from its serving BS which in turn may specify, in terms of time interval and metric, how the SS should report the scanning result back to the BS. Apart from SS-initiated scanning, the reservation of scanning intervals may be transmitted unsolicited by the BS. Based on the feedback from the SS, the BS may build a neighborhood list which is periodically broadcasted. The latter includes information for each neighbor BS regarding up- and downlink channel slot assignments. This consists of BS identifier and PHY synchronization field, thus, these parameters have not to be obtained while switching from one BS to another , which reduces handover latency.

In order to establish a link layer connectivity with the new BS, the SS has to convey information like its MAC address and capability information to the target BS. Besides, it has to go through the ranging process. In the traditional way, the SS "associates without coordination", i.e., it exchanges these information directly with the target BS over the wireless link. In the second mode, the serving BS coordinates the association

by forwarding these information to the target BS via the backbone allowing the SS to immediately start the ranging process in order to adjust its transmission power correctly. The third approach is "with network assistance". Additionally to mode two, target and serving BSs exchange the feedback of the ranging algorithm over the backbone and the serving BS provides a single, condensed answer to the SS. This scheme allows the SS to maintain multiple associations at a time reducing the latency of the handover process.

As part of optimizing the handover process, the exchange of security keys between BS and SS can also be shifted into the scanning phase prior to the handover. Even a direct communication between serving and target BS is foreseen neglecting the need for authorization via the wireless link during the handover. Additionally, a two-level hierarchy of keys similar to the specified method in 802.11r [10] is introduced. The authenticator function calculates BS-specific keys based on the MSK received from an authentication server which allows the BS to act as an authenticator relay [17], [18].

In order to provide a seamless handover support even for higher OSI-layers, IEEE 802.16e also amends optional support for fast handover, namely macro diversity handover and fast BS transition. Both cases require strict time synchronization of involved BSs including exchange of MAC state information as well as their operation on the same frequency. For macro diversity handover, involved BSs synchronously transmit downlink data such that diversity combining can be performed by the SS. For the uplink, traffic is received by all involved BSs such that selection diversity can be performed. The information on the up- and downlink slot assignment may be either conveyed by all BSs forming the diversity set or only by a single BS, the so-called anchor. For the fast BS transition approach, only a single BS anchor provides up- and downlink capacity. The continuous monitoring of the BSs' signal levels allows adding and dropping BSs as well as the decision on when to switch to a new anchor. [17]

C. IEEE 802.21

The IEEE 802.21 working group focuses on media independent handover services; the first draft version [19] was finished in March 2006. The goal is to optimize handovers between heterogeneous access technologies such that on-going services of end users are not terminated, i.e., services can be continued although a handover takes place. IEEE 802.21 covers wired as well as wireless technologies including media specifications of the IEEE 802 group as well as of 3GPP and 3GPP2.

IEEE 802.21 will discover and provide relevant pieces of information for handovers decisions to upper layers. This includes signaling of information about QoS support of access networks, network discovery, and network selection. In other words, the group will provide a framework for generic link layer purposes. Handover policies and handover decision entities are thereby out-of-scope.

For the generic link layer instance, IEEE 802.21 introduces a Media Independent Handover (MIH) function between layer-2 and upper layers. The MIH function will define generic SAPs and primitives to higher as well as to lower layers. This may

---

[1]IEEE 802.11F has expired. Its withdrawal has been voted on by the IEEE 802.11 working group in November 2005 and was approved by the IEEE SA in March 2006.
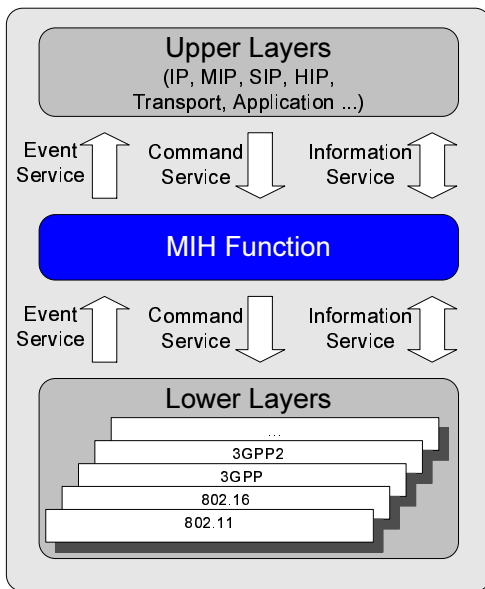
Fig. 1.   Placement and services of the IEEE 802.21 MIH function

later on require an adaptation of technology-specific SAPs of IEEE 802.11, IEEE 802.16, 3GPP, and 3GPP2.

Figure 1 shows the placement of the MIH function and its services. The MIH function is a logical entity which resides on MN as well as on network side. Pieces of information can be exchanged either locally within node's protocol stack by triggers or between MN and an access network entity via MIH-specific messages; for the latter part, IEEE 802.21 specifies the MIH Protocol.

The MIH function provides three services: Media Independent Event, Command, and Information Service. They are responsible for signaling state changes occurring at lower layers, coordination and control by higher layers, and information provision about the current and neighbor access networks, respectively.

IV. IETF

Historically, the work of the IETF towards mobility support originates in solving the well-known problem of *IP semantic overloading*: A mobile's IP address is used for routing purposes, i.e., it represents the node's network attachment point (NAP), as well as for identifying the mobile itself. Thus, a change of the IP address breaks established transport protocol connections as they traditionally employ the former for identifying involved endpoints.

Work approaching this problem can be classified in an *end-to-end-based mobility support* yielding to *sophisticated mobile IP schemes* and *auxiliary network enhancements for mobility support*. The former traditionally did not rely on sophisticated network features supporting mobility but in its evolution induced certain entities into the access network to decrease the signaling overhead and latency due to mobile end-points. The latter provides additional support from the access network regarding rerouting and context transfer, NAP detection, management of handover domains, or multihoming support.

A. End-to-End Approaches

Certain schemes handle mobility at the involved end-points: SCTP, HIP, and MIP try to solve mobility issues without any or with only minimal network support.

The base Host Identity Protocol (HIP) separates location and host identification by introducing host identifiers between the transport and the network layer. Instead of being bound to an IP address, higher layers use a representation of this host identifier — denoted as HIT (Host Identity Tag) which is a cryptographic hash over the host identifier — for their end-point addressing [20]. An enhancement of base HIP [21] supports changes of a single IP address as well as a mode with multiple IP addresses which allows mobility handling and multihoming, respectively. With signaling messages including a "locator" parameter, a host is able to inform its peer about other IP-address(es) under which it is reachable. For mobility handling, this option can be used to update a peer after an IP (sub)-net change.

The Stream Control Transmission Protocol (SCTP) was standardized by the IETF Signaling Transport (SIGTRAN) working group and was initially designed for the transport of telephone signaling data over IP networks [22]. It is a reliable transport protocol that is able to support multi-homing and multi-streaming within one connection [22], which are actually the main differences to TCP. With the optional Dynamic Address Reconfiguration (DAR) enhancement [23], which allows a reconfiguration of IP address(es) during an active communication session, SCTP is applicable for mobility handling on an end-to-end basis.

MIPv6 [24] has been designed to keep a single, permanent home address for a Mobile Node (MN) in case of nomadic movements between different IP subnets , whereby MT is reachable via its Care-of-Address (CoA) in foreign access networks. For true MNs which move during an on-going session between different subnets, MIPv6 is not sufficient: It requires that link level establishment has been completed prior any Layer 3 actions such as movement detection, Care-of-Address (CoA) configuration, and signaling between involved MIPv6 entities like Corresponding Node (CN), Home Agent (HA), and MN.

All end-to-end based approaches have the advantage that their is only low complexity required within the underlying networks, since all or most functionality is provided at the end points. However, severe disadvantages are high handover latencies as well as packet losses due to end-to-end signaling.

B. Sophisticated MIP Schemes

In order to enable movements between IP subnets during an on-going session, the IETF developed two RFCs in the *MIPv6 Signaling and Handoff Optimization (mipshop)* group — Fast MIPv6 (FMIPv6) and Hierarchical MIPv6 (HMIPv6).

FMIPv6 [25] targets to decrease packet losses by introducing a tunnel between Previous CoA (PCoA) and New CoA (NCoA). When a HO takes place, the Previous Access Router (PAR), which resides in the old subnet, forwards packets to NCoA. The New AR (NAR), which is located in the new subnet, buffers these packets and forwards them to MN after its

arrival. In order to be able to establish the tunnel between PCoA and NCoA, FMIPv6 assumes that either the MN or the PAR has a-priori knowledge about the NCoA.

The goal of HMIPv6 [26] is to reduce signaling overhead between the MN and its CNs or its HA, respectively. Therefore, it introduces a node with HA functionality — the Mobility Anchor point (MAP) — which can be located elsewhere in the hierarchy of routers. HMIPv6 introduces two CoA's: one to address MAP's subnet, the second to address the current location of MN within MAP's subnet. Thus, the MN sends binding updates only to MAP, since the outer CoA seen by CN remains. This reduces the binding update latency and thus also the handover delay, especially if distance between HA and MN / CN and MN is large.

Combinations of FMIP and HMIP have been proposed in [27] (expired in April 2006) and the Appendix of [26]. A simple integration of FMIP into HMIP results in a tunnel between PAR and NAR, which in turn has a high signaling overhead but further reduces handover latency and packet losses. In F-HMIP [27], the MAP sets up a tunnel with NAR, which then caches packets and forwards them after MN's registration. Another option includes the bi-casting of packets, i.e., MAP sends packets to PAR as well as NAR during the handover process.

Additionally, F-HMIP proposes a tunnel between MAPs in case that a MN performance a HO between subnets of different MAPs.

C. Auxiliary Enhancements for Mobility Support

Apart from end-to-end based mobility schemes and evolved sophisticated MIP derivatives, IETF working groups focus on mobility schemes increasingly providing mobility enhancements by the network, either in terms of context transfer and router discovery (SeaMoby), detecting IP network attachment points (DNA), network based mobility management (NETLMM), hierarchical key level for an enhanced and accelerated establishment of secure associations (MOBIKE and HoKEY), as well as flow-based mobility support using multiple available interfaces (MONAMI).

1) SeaMoby: The Seamless Mobility (SeaMoby) working group focused on context transfer, candidate access router discovery, and IP-paging.

The Context Transfer Protocol (CXTP) provides a generic framework to transfer flow- or service-specific information from one AR to another. Either the network or the mobile can initiate the transfer predictively, i.e., prior to the handover, or afterwards, which may still be beneficial as context re-establishment might take longer than the actual transfer. The only context defined so far is for IPv6 Multicast but other, even layer-2 specific context may be exchanged. [28], [29]

The Candidate Access Router Discovery (CARD) protocol allows to discover the identity, i.e., IP address, and capabilities of candidate access routers before conducting IP-level handover. Usually, scanning at link level results in finding a set of potential handover APs being associated with different or identical access routers (ARs) whose layer-2 identifier is used to request their IP address and capability information via the current AR. The resolution of candidate APs' L2 IDs to the IP address of their associated ARs is, e.g., achieved by a L2-L3 address mapping table manually configured and stored at each AR. [30] In practice, this limits CARD's usage to one administrative domain.

Work regarding IP-paging discontinued after publishing the problem statement and requirements. [31], [32]

2) DNA: The Detecting Network Attachment in IPv6 working group (DNA) targets a fast and efficient mechanism to detect IPv6 network attachment. It hereby assumes that lower layer services—e.g., IEEE 802.21 (ref. to Section III-C)—indicate "link-up" and "link-down" events. [33] Upon the reception of those, nodes may send router solicitations (RSs) to determine the IP subnets available on the new link. Two means to reduce the size of router advertisements (RAs) are discussed: the *landmark option* and common router *identifier prefix* usage. In the former case, a node includes a shortened routing prefix in its RS querying for routers having registered to this prefix; whereas in the latter case, several routers on link commit to common, short "link identifiers" which can be used in unsolicited router advertisements. [34] Also, DNA discusses to make NAPs, e.g., an IEEE 802.11 AP, caching RAs and send them immediately upon reception of a link-layer establishment request. [35]

3) NETLMM: The Network-based Localized Mobility Management (NETLMM) working group fills the gap between layer-2-based mobility schemes as can be found in Wireless LANs and IP-based macro mobility schemes. The latter suffer from a number of problems when used in smaller coverage areas: If the mobility anchor point used in macro mobility schemes is far away from the mobile node's NAP, location updates incorporate large *update latencies*. Additionally, each movement from one NAP to another requires a location update yielding to a large *signaling overhead* which at the same time reveals the mobile node's topological location (*location privacy*). [36] A number of proposals for realizing a NETLMM-based mobility scheme are discussed (see [37]–[43]).

Proposed solutions generally introduce two functional elements: the mobility anchor points and the mobile access gateways. A mobility anchor point (MAP) — denoted as edge MAP or local MAP—acts as a light location server and "entry point" to the considered mobility domain. The local access routers—also denoted as mobile access gateways (MAG)—terminate link-layer specific mobility support. [37], [43]

Since the NETLMM working group condemns use of HMIPv6 due to its prerequisite to be included in the mobile node's IP stack, Raman et al. propose using Proxy Mobile IP [44]–[46]. It does not require the modification of the mobile node's IP stack but shifts the functionality of the MIP client into the network. Both, MIPv6 [24] and HMIPv6 [26], either one acting as a local MAP, are considered. [38], [42], [44]

4) MOBIKE: The IKEv2 Mobility and Multihoming (MOBIKE) working group was formed to enhance the Internet Key Exchange Protocol (IKEv2) [47] in order to support roaming, mobility, and multihoming. This enhancement became necessary as an IKE security association, which mutually authenticates two hosts, employs the latter's IP addresses to identify the secure association. Thus, changing the end-point's

IP address requires a full re-keying. [48] The standard track protocol addresses this problem and additionally enables a limited support of multi-homed nodes. [49]

5) HOKey: Handover latencies are affected significantly by authentication mechanisms that control network access. The IETF Handover Keying (HOKey) working group considers extensions to the actual IETF Extensible Authentication Protocol (EAP) [9] framework in order to mitigate handover delays caused by authentication exchanges: securing context transfers among access policy enforcement points is mandatory when pre-authentication schemes are not available in the underlying layer-2 technology. The group discusses fast re-authentication schemes avoiding full-featured EAP exchanges during a handover. Finally, the working group focuses on means to allow an EAP/AAA server in the visited domain to handle an authentication independent from the home network's AAA services. This resembles mechanisms from existing 2.5G/3G networks, where HLR and VLR exchange triplets for local authentication purposes. [18], [50], [51]

6) MONAMI: The Mobile Nodes And Multiple Interfaces in IPv6 (MONAMI6) working group was formed recently. It tackles the issues of interface selection, concurrent use of multiple CoAs, simultaneous location in Home and Foreign Networks, as well as flow redirection. [52]–[55] Thus, the group follows an application- / flow-based mobility management scheme which rises the question which granularity of a mobility framework should be available in a future IP-based mobility-aware network.

V. 3GPP

One of the requirements on the 3GPP system was native support of QoS. 3GPP therefore focussed on make-before-break handovers and devised its own protocol, the GPRS Tunnelling Protocol (GTP) [56] to handle mobility.

Requirements for future mobile telecommunication networks, beyond the current 3GPP system, were formulated by the ITU-T [57]. They include the integration of heterogeneous access technologies and the seamless handover between different access technologies, which were already in the Introduction of this paper as common traits of ongoing standardization efforts. The ITU-T requirements however also include functionality such as paging support and context transfer. Furthermore, interworking with established AAA and security schemes as well as support for location privacy is necessary.

3GPP is now actively working on its evolution, basically in-line with the ITU-T requirements. According to the requirements formulated in [58], [59], mobility and service continuity between heterogeneous access systems shall be supported. In the System Architecture Evolution (SAE) work item it is being discussed to what extent an IP-based solution may serve as a basic building block to satisfy these requirements. The differences in the service model between 3GGP and IP-based networks — selling high-quality user services rather than just connectivity which translates also in network-controlled mobility — makes it however questionable whether the 3G core network can employ an IP-based mobility framework.
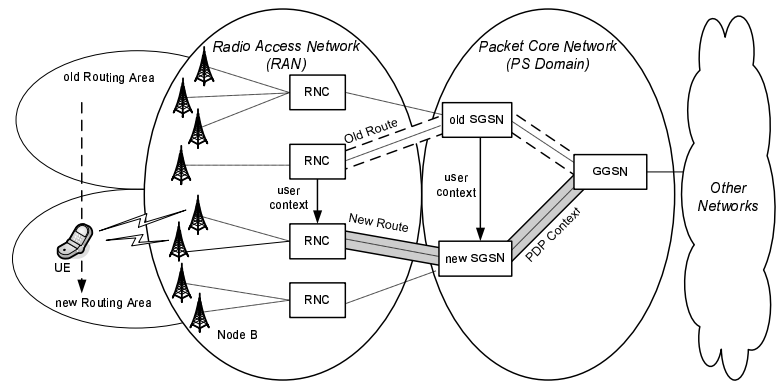

Fig. 2. Architecture of a 3GPP network with a User Equipment (UE) performing a handover from route (1) to route (2).

This section starts with a discussion on how operator-controlled, seamless mobility has been achieved in 3GPP networks so far. Afterwards, it explains the different options for evolution of mobility support currently explored in the work item System Architecture Evolution in 3GPP.

A. Mobility Support in 3GPP Networks of Release 99

Mobility support in the original 3GPP networks (Release 99, published in 2000) is tightly coupled to the hierarchical architecture illustrated in Fig. 2 [56]. The User Equipment (UE) sends its data packets via radio access points (Node B's) to a Radio Network Controller (RNC). In addition to forwarding data packets, the RNC controls the Node Bs as well as the mobility of UEs which are currently having an active session ("Connected Mode"). The RNC forwards the data packets to a Serving GPRS Support Node (SGSN) in the Core Network. In addition to forwarding data packets, the SGSN controls the mobility of the UE also when it is not engaged in a session ("Idle Mode"), and is responsible for other control functions such as security. Data packets finally pass through the Gateway GPRS Support Node (GGSN) and from there to the destination. This destination can be in the same 3GPP network or in other networks, including the Internet.

A UE in Connected Mode has a Serving RNC assigned to it. From the Serving RNC via SGSN to GGSN a tunnel is established for the UE's data packets with the GTP protocol. This tunnel is known as PDP context. A specific QoS is associated with this tunnel. When the UE moves in Connected Mode, the Serving RNC may decide to initiate a handover between Node B's, based on the radio conditions communicated by both UE and Node B. In order to achieve make-before-break, the Serving RNC initiates resource reservation along the new section of the path and, once everything is in place, it tells the UE to perform the handover. In fact, macrodiversity can be employed, i.e. the UE may be connected to more than one Node B simultaneously and communicate in a multi-path fashion. The Serving RNC then adds and removes Node Bs as appropriate. This form of sliding handover is called "soft handover". When the UE moves very far and is connected to Node B's that are no longer controlled by the Serving RNC, the Serving RNC may be relocated to another RNC. User

context such as QoS is transferred from old to new Serving RNC, and the PDP context is moved. The process of Serving RNC relocation is independent of the process of handover and may be performed any time.

In Idle Mode, the UE has no Serving RNC. The UE is only attached to a specific SGSN. Each SGSN is assigned a specific set of Routing Areas. The UE listens to cell broadcasts on the local Routing Area and informs the SGSN about its location on a regular basis. When a session request comes in for the UE via the GGSN, the GGSN first finds the right SGSN by contacting the 3GPP networks central database, and the SGSN then sends a paging request to all cells in the UEs current Routing Area. Upon reception of the paging request, the UE switches into Connected Mode and obtains a Serving RNC.

When the UE moves away from the location where it originally booked into the network it may reach a Routing Area which is not controlled by the SGSN it is attached to. In this case it attaches to a new SGSN. User context is transferred from old SGSN to new SGSN; also the GTP tunnel is relocated.

To summarize, in a 3GPP network, the UE is responsible for reporting radio conditions and its location to the network. In the network, RNCs and SGSNs collaborate to control and perform a seamless handover.

B. Mobility in future 3GPP Networks

Already in Release 6, published in 2005, it was specified how a 3GPP subscriber can achieve access to the 3GPP network via a WLAN. The WLAN Access Network is connected to the 3GPP network via a Packed Data Gateway (PDG). Authentication, authorization, and charging thereby is performed in the 3GPP network. Requests from UEs are forwarded by the WLAN Access Network to the 3GPP network [60]. However, a handover between WLAN and the 3GPP Radio Access Network is not possible.

Also Release 6 specified how a user can more generally roam between cellular networks, public and private unlicensed wireless networks (e.g. 802.11 networks located at user premises), or wired networks. The corresponding specification [61] is based on previous work by the UMA (Unlicensed Mobile Access) project. Whereas for 3GPP-WLAN interworking described above, the UE employs IEEE and IETF protocols only for corresponding with the WLAN AP, here the UE employs 3GPP specific protocols for corresponding with the UMA AP. Mobility support is also based on 3GPP specific protocols.

The next release of the 3GPP network specification currently discussed by 3GPP in the context of SAE is expected to introduce a major update regarding architecture, protocols, and radio technology. The goal is on the one hand to considerably increase radio interface bandwidth. The goal is also to support access to the 3GPP network via multiple non-3GPP access networks, including e.g. WiMAX, and to support handover between these access networks and a 3GPP network. At this point it is not foreseen that this handover be seamless.

Fig. 3 shows the current status of the architecture debate [59]. An evolved 3GPP RAN is connected to an evolved Packet
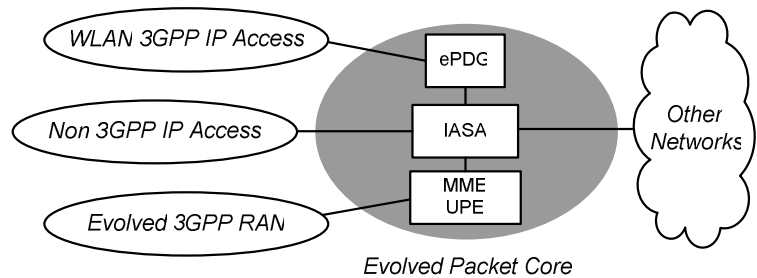

Fig. 3. Architecture for an evolved 3GPP network

Core. Roughly, the former SGSN is now subdivided into an entity handling control functions, the Mobility Management Entity (MME), and an entity handling the user traffic, the User Plane Entity (UPE). The final functionality split between the evolved RAN and the evolved Packet Core is not yet clear. MME and UPE may be split and the MME, e.g., moved into the RAN. In this case, the MME is combined with the "evolved RNC" such that one hierarchy level of the control architecture is removed. In any event, the MME is responsible for intra-3GPP mobility control. Non-3GPP access systems are connected to an "evolved GGSN" called Inter Access System Anchor (IASA). A WLAN access system continues to be connected to an (evolved) PDG. The current 3GPP RAN and PS Domain shown in Fig. 3 are deployed in parallel to the evolved Packet Core and evolved 3GPP RAN.

For supporting mobility, particularly between the 3GPP RAN and non-3GPP access technologies, several options are being debated [59]

Regarding mobility within the 3GPP system, i.e. within the evolved Packet Core and towards the (legacy) Packet Core, GTP is maintained. However, for mobility between non-3GPP access networks and the evolved Packet Core, Mobile IP is employed. Thereby, both IPv4 and IPv6 shall be supported. Mobile IP usually runs between the mobile node, i.e. UE, and the home agent, i.e. the IASA. Hence the UE becomes involved in mobility control, which on the one hand takes away control from the network, and furthermore implies existing UEs must be updated. An alternative is the usage of Proxy MIP [38] which is however still in very early draft state. In addition to the global mobility supported by Mobile IP, usage of a micro-mobility protocol such as one of the protocols developed by the IETF NETLMM Working Group [37] is being discussed, which also allows network-based mobility control.

VI. CONCLUSION

Standardization bodies move towards a framework for seamless mobility support involving heterogeneous access networks. The IEEE has provided enhanced layer-2 mechanisms which facilitate seamless handover—at least supporting QoS constrains for VoIP—and works on a media independent handover framework to allow upper layer protocols to probe and control the underlying link regardless of its technology.

IETF has invested considerable effort in auxiliary services that are required to allow true ubiquitous mobile access.

This includes IP subnet detection, host alerting, multi-homed operation, context transfer, candidate router selection, and soft state during handover. Further more, combined F-HMIP approaches aim to reduce handover latencies as well as packet loss. With the NETLMM framework, development of micro-mobility protocols has regained interest in IETF.

In the SAE work of 3GPP, IP-based mobility mechanisms are under consideration in order to achieve handover to heterogeneous access networks.

In summary, current and emerging paradigms enabling seamless handover cover almost all layers of the OSI protocol stack. In order to provide neglectfully small handover delays for real-time traffic, e.g. VoIP or even telemetry applications, sophisticated layer-2-based approaches are essential. Only in cases where this mobility support is not sufficient — due to higher layer's address reconfiguration, context transfer, tunneling etc. — other mobility schemes at layer-3 or above are required. Solutions for the latter case always have to rely on a fast and efficient layer-2 connection (re-)establishment. For hosts which are equipped with more than one network interface card, higher layer solutions like HIP or SCTP may be sufficient due to their multi-homing capabilities without any layer-3 mobility support . However, this requires sufficient knowledge about the handover decision such that signaling and bi-casting of user data can be evoked timely. It remains questionable whether this timeliness is feasible without knowing anything about underlying wireless access technologies. Due to the error-prone nature and high non-predictiveness of wireless channels, isolated approaches are rather bound to fail for true seamless mobility. Thus, mobility management schemes are most advantageous with a cross-layer approach, i.e, when they consider knowledge of other involved layers by using IEEE 802.21 for their information exchange.

Standardization bodies have clearly started to work on mobility support on all layers and are evolving to consider cross-layer-based approaches — or provide, at least, interfaces between different layers. Nevertheless, all standards only provide mechanisms for mobility support and do not tackle the interworking between different operators: integrating or merging policies to gain and seamlessly maintain network access while on the move is still not covered and is expected to remain a future challenge.

REFERENCES

[1] A. Dutta, S. Das, T. Chiba, H. Yokota, A. Idoue, K. D. Wong, and H. Schulzrinne, "Comparative analysis of network layer and application layer ip mobility protocols for ipv6 networks," in *Wireless Personal Multimedia Communications (WPMC)*, September 2006, pp. 6 – 10.
[2] M. Emmelmann, B. Rathke, and A. Wolisz, *Wireless Metropolitan Area Networks: WiMAX and Beyond*. Auerbach Publications, CRC Press, 2006, ch. Mobility Support for Wireless PAN, LAN, and MAN, iSBN: 0849326249.
[3] A. R. Prasad, "The futre re-visited," *Wireless Personal Communications, Int. Journal on,*, vol. 37, May 2006.
[4] M. Emmelmann, "Influence of velocity on the handover delay associated with a radio-signal -measurement-based handover decision," in *Proc.of IEEE 62nd Vehicular Technology Conference Fall 2005*. Dallas, TX, USA: IEEE, Sept. 2005.
[5] S. Mohanty and I. Akyildiz, "A cross-layer (layer 2 + 3) handoff management protocol for next-generation wireless systems," *Mobile Computing, IEEE Transactions on*, vol. 5, no. 10, pp. 1347–1347, October 2006.
[6] J. Spilman, "Test methodology for measuring bss transition time," IEEE 802.11 TGr Fast Base Station Transition, Working Document 802.11-04/0748r1, July 2004.
[7] H. Velayos and G. Karlsson, "Techiques to reduce ieee 802.11b mac layer handoff time," Royal Institute of Technology, Laboratory for Communication Networks Department of Microelectronics and Information Technology Stockholm Sweeden., Tech. Rep., April 2003.
[8] IEEE 802.11k, "Ieee 802.11k/d4.0 – radio resource measurement, amandment 9 to 802.11-revma-d5.2," IEEE, March 2006.
[9] B. Aboba, L. Blunk, J. Vollbrecht, J. Carlson, and H. Levkowetz, "Extensible authentication protocol (eap)," IETF, RFC 3748, June 2004.
[10] IEEE 802.11r, "Ieee 802.11r — fast bss transition, amandment 2 to 802.11-revma," IEEE, March 2006.
[11] E. Qi, S. Bangolae, K. Sood, and J. Walker, "Bss transition optimizations and analysis in wlan," in *Wireless Personal Multimedia Communications (WPMC)*, September 2006, pp. 11 – 15.
[12] IEEE 802.11F, "Ieee 802.11f — trial-use recommended practice for multi-vendor access point interoperability via an inter-access point protocol across distribution systems supporting ieee 802.11," IEEE, June 12 2003. [Online]. Available: http://standards.ieee.org/getieee802/download/802.11F-2003.pdf
[13] V. Anantha and R. Skidmore, "Need for site specific / location specific rf management," IEEE 802.11 TGv Network Management, Working Document 802.11-05/740r0, July 2005.
[14] E. H. Qi, J. Walker, M. Rudolf, J. Kwak, and J. Epstein, "Proposal for load balancing," IEEE 802.11 TGv Network Management, Working Document 802.11-05/1065r2, March 2006.
[15] B. Wang, F. Li, and J. Li, "Solve frequent handover requirement," IEEE 802.11 TGv Network Management, Working Document 802.11-05/1071r3, November 2005.
[16] B. Wang, F. Li, J. Li, and T. Weong, "Frequent handover," IEEE 802.11 TGv Network Management, Working Document 802.11-06/1280r2, January 2006.
[17] IEEE 802.16, "Ieee 802.16 — air interface for fixed and mobile broadband wireless access systems amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands," IEEE, February 2006, iEEE Std. 802.16e-2005 and 802.16-2004/Cor1-2005.
[18] M. Nakhjiri, M. Parthasarathy, J. Bournelle, H. Tschofening, and R. M. Lopez, "Aaa based keying for wireless handovers: Problem statement: draft-nakhjiri-aaa-hokey-ps-03," IETF draft, June 2006.
[19] IEEE 802.21, "Ieee 802.21/d1.0 — draft ieee standard for local and metropolitan area networks: Media independent handover services," IEEE, March 2006.
[20] R. Moskowitz, P. Nikander, P. Jokela, and T. Henderson, "Host identity protocol," IETF, draft, June 2006.
[21] T. Henderson, "End-host mobility and multihoming with the host identity protocol," IETF, draft, June 2006.
[22] R. Stewart, Q. Xie, K. Morneault, C. Sharp, H. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang, and V. Paxson, "Stream control transmission protocol," IETF Standards Track, RFC 2960, October 2000.
[23] R. Stewart, Q. Xie, M. Tuexen, S. Maruyama, and M. Kozuka, "Stream control transmission protocol (sctp) dynamic address reconfiguration," IETF, draft, October 2006.
[24] D. Johnson, C. Perkins, and J. Arkko, "Mobility support in ipv6," IETF, RFC 3775, June 2004.
[25] R. Koodli, "Fast handovers for mobile ipv6," IETF, RFC 4068, July 2005.
[26] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, "Hierarchical mobile ipv6 mobility management (hmipv6)," IETF, RFC 4140, August 2005.
[27] H. Jung, H. Soliman, and S. Koh, "Fast handover for hierarchical mipv6 (f-hmipv6)," IETF, draft, October 2005.
[28] J. Loughney, M. Nakhjiri, C. Perkins, and R. Koodli, "Context transfer protocol (cxtp)," IETF, RFC 4067, July 2005.
[29] J. Kempf, "Instructions for seamoby and experimental mobility protocol iana allocations," IETF, RFC 4065, July 2005.
[30] M. Liebsch, A. Singh, H. Chaskar, D. Funato, and E. Shim, "Candidate access router discovery (card)," IETF, RFC 4066, July 2005.
[31] J. Kempf, "Dormant mode host alerting ("ip paging") problem statement," IETF, RFC 3132, June 2001.
[32] J. Kempf, C. Castelluccia, P. Mutaf, N. Nakajima, Y. Ohba, R. Ramjee, Y. Saifullah, B. Sarikaya, and X. Xu, "Dormant mode host alerting ("ip paging") problem statement," IETF, RFC 3154, August 2001.
[33] J. Choi and G. Daley, "Goals of detecting network attachment in ipv6," IETF, RFC 4135, August 2005.

[34] S. Narayanan, "Detecting network attachment in ipv6 networks (dnav6): draft-ietf-dna-protocol-03.txt," IETF draft, October 2006.

[35] J. Choi, D. Shin, and W. Haddad, "Fast router discovery with l2 support: draft-ietf-dna-frd-02.txt," IETF draft, August 2006.

[36] J. Kempf, K. Leung, P. Roberts, K. Nishida, G. Giaretta, and M. Liebsch, "Problem statement for ip local mobility: draft-kempf-netlmm-nohost-ps-05.txt," IETF draft, September 2006.

[37] I. Akiyoshi and M. Liebsch, "Netlmm protocol: draft-akiyoshi-netlmm-protocol-00.txt," IETF draft, October 2005.

[38] S. Gundavelli and K. Leung, "Localized mobility management using proxy mobile ipv6: draft-gundavelli-netlmm-mip6-proxy-00.txt," IETF draft, November 2005.

[39] J. Wood and K. Nishida, "Edge mobility protocol (emp): draft-wood-netlmm-emp-base-00.txt," IETF draft, October 2005.

[40] G. Giaretta, I. Guardini, and E. Demaria, "Network-based localized mobility management (netlmm) with distributed anchor routers: draft-giaretta-netlmm-protocol-00.txt," IETF draft, October 2005.

[41] M.Parthasarathy, B. Patil, and R. Koodli, "Network-based fast handovers for local mobility (nflm)," IETF draft, October 2005.

[42] V. Raman, A. Bedekar, A. ingh, and S. Kalyanasundaram, "A protocol for network-based localized mobility management: draft-raman-netlmm-protocol-00.txt," IETF draft, February 2006.

[43] H. Wang, Y. Zhang, and Y. Xia, "Netlmm protocol: draft-wanghui-netlmm-protocol-00.txt," IETF draft, April 2006.

[44] S. Gundavelli, K. Leung, and V. Devarapalli, "Proxy mobile ipv6: draft-sgundave-mipv6-proxymipv6-00," IETF draft, October 2006.

[45] D. Premec and D. Damic, "Mobility management for ipv6 hosts using proxy mobile ipv4: draft-premec-mip4-ip6-proxy-mip4-00.txt," IETF draft, June 2006.

[46] K. Leung, G. Dommety, P. Yegani, and K. Chowdhury, "Mobility management using proxy mobile ipv4: draft-leung-mip4-proxy-mode-01.txt," IETF draft, June 2006.

[47] C. Kaufman, "Internet key exchange (ikev2) protocol," IETF, RFC 4306, December 2005.

[48] T. Kivinen and H. Tschofening, "Design of the ikev2 mobility and multihoming (mobike) protocol," IETF, RFC 4621, August 2006.

[49] P. Eronen, "Ikev2 mobility and multihoming protocol (mobike)," IETF, RFC 4555, June 2006.

[50] Z. Cao, Y. Ma, H. Deng, and Q. Li, "Handover key hierarchy based on extended master session key (emsk) derivation: draft-cao-hoakey-hierarchical-hokey-00.txt," IETF draft, June 2006.

[51] M. Nakhjiri, "A keying hierarchy for managing wireless handover security: draft-nakhjiri-hokey-hierarchy-02," IETF draft, June 2006.

[52] R. Rakikawa, T. Ernst, and K. Nagami, "Multiple care-of addresses registration: draft-ietf-monami6-multiplecoa-01.txt," IETF draft, October 2006.

[53] R. Kuntz, M. Tsukada, E. Paik, T. Ernst, and K. Mitsuya, "Evaluating multiple mobile routers and multiple nemo-prefixes in nemo basic support: draft-kuntz-nemo-multihoming-test-02.txt," IETF draft, July 2005.

[54] T. Ernst, N. Montavont, R. Wakikawa, C. Ng, and K. Kuladinithi, "Motivations and scenarios for using multiple interfaces and global addresses: draft-ietf-monami6-multihoming-motivation-scenario-01.txt," IETF draft, October 2006.

[55] N. Montavont, R. Wakikawa, T. Erns, C. Ng, and K. Kuladinithi, "Analysis of multihoming in mobile ipv6: draft-ietf-monami6-mipv6-analysis-01.txt," IETF draft, June 2006.

[56] 3G-Partnership-Project, "3gpp ts 23.060 v7.0.0 (2006-03): 3rd generation partnership project; technical specification group services and system aspects; general packet radio service (gprs); service description; stage 2 (release 7)," 3GPP Technical Specification, 2006. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/23060.htm

[57] ITU-T, "Itu-t q supplement 52: Nni mobility management requirements for systems beyond imt-2000," December 2004.

[58] 3G-Partnership-Project, "3gpp ts 22.278 v1.0.0 (2006-06): 3rd generation partnership project; technical specification group services and system aspects; general packet radio service (gprs); service requirements for evolution of the 3gpp system (release 8)," 3GPP Technical Specification, 2006. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/22278.htm

[59] ——, "3gpp tr 23.882 v0.10.0 (2006-01): 3rd generation partnership project; technical specification group services and system aspects; 3gpp system architecture evolution: Report on technical options and conclusions (release 7)," 3GPP Technical Specification, 2006.

[60] ——, "3gpp ts 23.234 v7.1.0 (2006-03): 3rd generation partnership project; technical specification group services and system aspects; 3gpp system to wireless local area network (wlan) interworking; system description (release 7)," 3GPP Technical Specification, 2006.

[61] ——, "3gpp tr 43.318: 3rd generation partnership project; technical specification group services and system aspects; generic access to the a/gb interface gan; system description;(release 7)," 3GPP Technical Specification, 2006.

Marc Emmelmann (emmelmann@ieee.org) graduated summa cum laude from Technical University of Berlin holding a Master of Science degree in Computer Engineering. He is currently working for the Technical University of Berlin, faculty of computer science and electrical engineering. As a member of the Telecommunication and Networks Group (TKN), his current research focuses on wireless networks supporting seamless moblity. Special interest has evolved in MAC-layer supported seamless handover with QoS guarantees. From 1997 to 2002, Marc was affiliated with Fraunhofer Institut for Open Communication Systems (Fokus) where he was responsible for the analysis and simulation of transport protocols in a satellite based environment. Previously, he had joined LCB Systemhaus Berlin in a project targeted at the in-circuit-based software development for handhelds (PDAs). Mr. Emmelmann is an active member of the IEEE having served in several Technical Program Commitees as well as a reviewer for numerous conferences and journals. He is actively participating in the IEEE standardization process holding "voting membeship" status within the 802.11 Working Group.

Sven Wiethoelter (wiethoelter@ieee.org) received the Master (Dipl.-Ing.) degree in electrical engineering from the Technical University of Berlin (TUB), Germany, in 2005. Being a member of the the Telecommunication Networks Group (TKN) at TUB since 2002, his work was firstly focused on medium access protocols for WLANs and sensor networks. He is currently working towards his PhD degree, whereby his research interests include handovers, resource management, and mobility support in wireless networks.

Andreas Koepsel (koepsel@tkn.tu-berlin.de) received his Dipl.Ing. of Electrical Engineering in June 2000 at the Technical University of Berlin. After working as a research assistant in the field of audio transmission over IEEE 802.11 based Wireless LANs at Technical University of Berlin, he co-founded a startup company in the area of Wireless LAN security systems including development of security and OAM components. Currently he is engaged as a Ph.D. student with Siemens AG working in the area of new naming, addressing, and resolution services for information access.

Cornelia Kappler (cornelia.kappler@siemens.com) studied physics at Munich University, Harvard University, and the University of Toronto. In 1995 she received a Ph.D. from the University of Toronto. Later she switched fields into communication networks, working for NEC Networking Laboratories, Berlin, Germany, and, since 2000, for Siemens Networks in Berlin. She is a project manager of international research projects and actively contributes to standardization in the IETF and 3GPP. Her research interests focus on 4G networks and signaling protocols.

Adam Wolisz (awo@ieee.org) (Diploma in engineering, 1972, Doctoral Degree, 1976, Habilitation 1983 - Silesian University of Technology, Gliwice) works since 1980 on computer networks and distributed systems. He has been with Polish Academy of Sciences (until 1990), and later with the Research Institute GMD-Fokus in Berlin (1990-1993). Since 1993 he has joined the Technische UniversitŁt of Berlin (TUB) where he is chaired Professor for Telecommunication Networks and since 2001 Executive Director of the Institute for Telecommunication Systems. He has served as the Dean of the Faculty of Electrical Engineering and Computer Science in the period 2001-2003. Since Summer 2005 he is also Adjunct Professor at the Dept. EE&CS, University of California, Berkeley. His research interests are in architectures and protocols of communication networks. Recently he is focusing mainly on wireless/mobile networking and sensor networks.