Service Architecture for Infrastructure based Multi-hop Networks based on SIP

Bernhard H. Walke, Ian Herwono, Ralf Pabst
Communication Networks, Aachen University (RWTH)
Email: {walke | ian | pab}@comnets.rwth-aachen.de

Abstract—**A new service concept for a mobile broadband system called the Wireless Media System (WMS) is introduced. The WMS will provide broadband access to terminals with medium velocity of movement and is embedded into a cellular radio network to support a high velocity of terminals with medium transmission rate. An Intelligent Service Control and advanced radio resource management for a power-optimized exploitation of multiple air interfaces will provide a virtually continuous radio connectivity to the WMS although the physical radio connectivity might be discontinuous. Broadcast, multicast and single-cast services will be combined to minimize the number of transmissions needed to provide the contents requested by the users. The concept of the Intelligent Service Control as well as basic service concepts and multimedia applications and usage scenarios are presented.**

Keywords—Mobile Broadband System; Wireless Media System; Media Point; Intelligent Service Control; SIP

I. Introduction

The so-called Wireless Media System (WMS) [1, 2, 3, 4, 5, 6] is based on the integration of a cellular mobile broadband system for "hot area" coverage with a modern cellular mobile radio system for large area coverage. While aiming at an especially high spectrum efficiency the WMS will provide a new air interface with data rates of up to 100 Mbit/s at terminal speed up to 60 km/h. The radio part of the WMS is characterized with dynamic channel selection, link adaptation, power control, smart antennas and re-configurable terminals. The known candidate spectrum bands for operation of the new systems, e.g., beyond 3 GHz and most probably even beyond 5 GHz will allow the deployment of small ("invisible") base stations, called Media Point, Hamburger sized (including antenna) mounted below roof top (say on signal posts or in street lamps) with low EIRP value of up to 1 W outdoors and 200 mW indoors. To provide the radio coverage the Media Points (MP) operate either as Access Points to the core network or as Fixed Wireless Routers (FWR) in a pre-planned multi-hop communication based infrastructure. Smart antennas and beam forming at terminals and base stations (MP and FWR) are used to achieve higher spectrum efficiency and lower radio exposition of humans.

Any kind of communication services, e.g., voice, video, data, etc. with the appropriate service specific Quality of Service (QoS) parameters can be supported by WMS. The mobile terminals must be able to support both, mobile cellular radio and the WMS air interfaces, but not at the same time. Future mobile broadband systems are characterized by a high asymmetry of their data traffic with an especially high load on the downlink, compared to the uplink. The goal is to carry in densely populated areas most of the high and medium bit rate multimedia traffic by the MPs of the WMS. In addition, some amount of the other traffic classes (like voice) also can be carried by the WMS.

The discontinuous radio coverage available from the WMS will be hidden to a maximum degree to the users by an Intelligent Service Control (ISC) so that the subscribers are provided with the contents they need within given time limits. The ISC simulates a virtual continuous radio network connectivity to applications of the wireless terminals. New service concepts like the integration of combined push-pull services and of broadcasting are applied within the WMS service architecture.

II. Service Architecture

The core network of the Wireless Media System (WMS) applies an Intelligent Service Control (ISC) to ensure that data communicated over the various radio access networks is continuously made available to a Mobile Terminal (MT) even if the radio connectivity to the WMS is interrupted time-wise owing to the discontinuous radio coverage available from the WMS. The ISC simulates to user applications a continuous connectivity of the MT to the WMS by making extensive use of caching at MTs and MPs or their nearby controllers. Spontaneous access via the WMS to contents data is typically executed with a situation-specific delay, since the respective MT must wait until it has reached an area with WMS radio supply – optionally the service might be provided from the mobile radio system, immediately.

An MT associated to the WMS, when reaching an MP served area refers to the session already established earlier with its remote application, receives the data it had requested earlier at a very high data rate from the cache of the respective MP controller and caches it locally for later use. The quantity of data transmitted should be large enough to accommodate the expected duration of the local processing while a terminal might not be connected to the system, e.g., mailbox content, MP3 music file or other large data file. An MT transmits all data waiting for transmission to the MP as soon as it reaches its coverage area.

MTs can use all services known from cellular networks and the Internet, i.e., voice, data transmission, reception of broadcast transmissions, etc. and may operate interactive multimedia connections. The ISC uses the mobility management and localization function of the advanced 3G radio system and dynamically tunnels data across fixed networks. Handover between adjacent cells served by MPs is replaced by fast re-association to the next MP.

In Fig. 1 two sub-networks of a WMS are shown, each with its Media Point Controller (MPC) comprising the cache storage for all the MPs in the sub-network. The dotted curve relates to the route of a car that originally had used the mobile radio system, e.g., to transmit a request for large volume data and later is roaming through MP sub-networks where it crosses a number of areas served by MPs of the WMS. The ISC will provide as much data as possible at the downlink, according what was requested by the terminal in the car.

It is clear from Fig. 1 that multicast data could optimally be broadcast from the base stations of the hexagonal cells, namely via the cellular radio system. One other possibility is to realize a multicast based service in the WMS by caching the respective data in the MPCs and pushing a copy down to an MT as soon as it shows up at one of the related MPs. This approach is taken in the WMS and will allow a perfect selective multicast service in that only pico-cellular MP capacity is used to reach the MTs that have subscribed to a multicast service.

The ISC supervises and controls all of the sessions of any mobile terminal in the combined systems of cellular mobile radio and the WMS. Its purpose is to control in an optimal way how to provide communication services related data to mobile terminals. The ISC is a logical unit that co-ordinates and comprises the following functions:

- mobility management of all of the terminals,
- localization of the terminals in both, the mobile network and WMS,
- Subscriber Identity Module (SIM) common with cellular radio network,
- dynamic tunneling of data through core networks involved to serve a mobile terminal,
- authentication, authorization, accounting and encryption,
- caching of data for mobile users,
- self-learning algorithms to understand the behavior of a user and to take advantage of knowledge on the type of MT and applications running on it,
- Quality of Service (QoS) support depending on the used network.

The ISC may be a centralized or decentralized unit. It simulates a continuous broadband access service to mobile terminals although the radio network coverage of the WMS is discontinuous only.

III. Multimedia Applications And Usage Scenarios

Broadband wireless access for mobile (moving) terminals, e.g., for terminals installed in vehicles, poses higher system requirements than for stationary terminals. It has been found that multimedia (MM) services to mobile users in many application scenarios could be satisfied by a much more cost-effective network architecture than known from cellular mobile radio systems, which then basically serve as a back-up system to provide services like voice and data that must be brought instantaneously to users without any cost concern [3], [7]. This architecture is

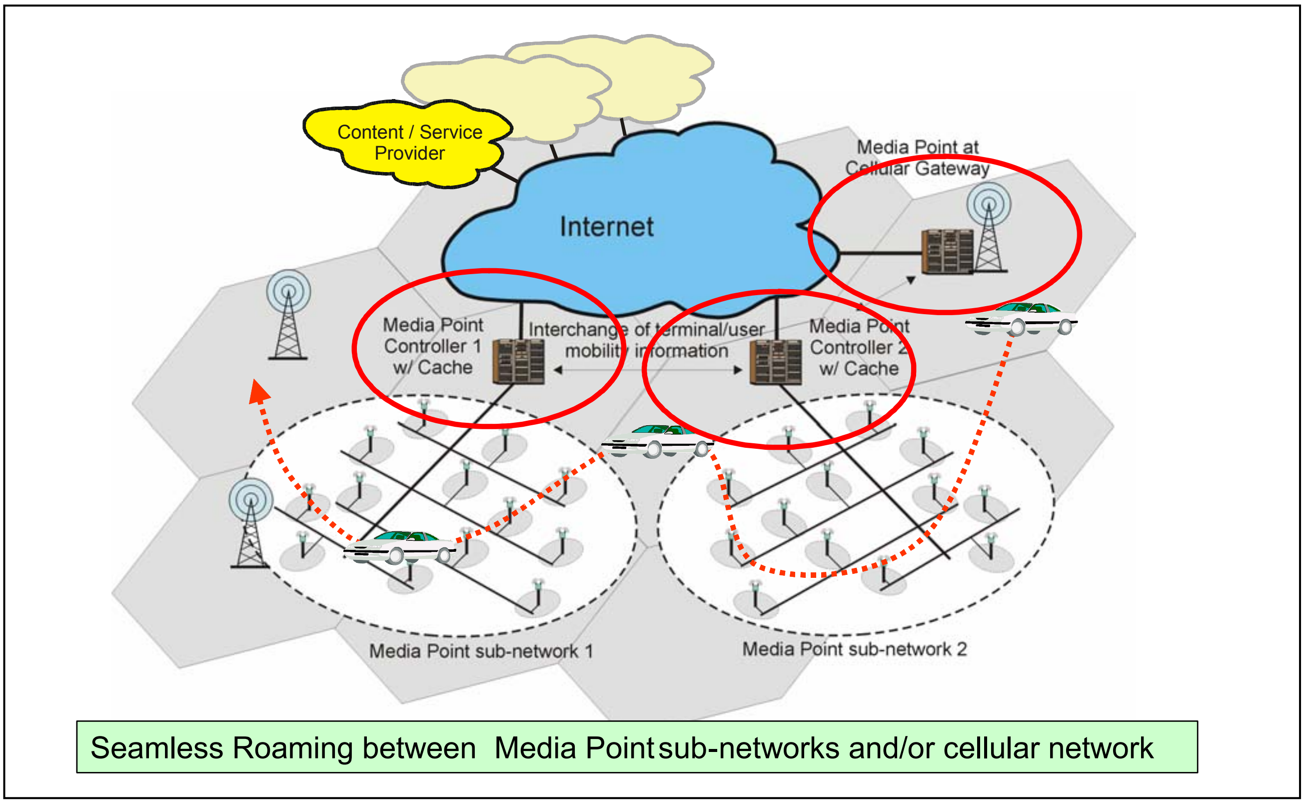

Figure 1. Media Point Controllers (as part of the ISC) to cache broadband data in preparation for download to mobile terminals

realized by the WMS approach. Only a perfect combination between a very high bit-rate wireless communication system and an ISC in terminal and network, that takes the terminal, application and network specific data into account to provide multimedia contents to mobile users at specific wireless access points, would lead to a success of the WMS.

A. Basic Service Concepts

The WMS is designed to realize the concept of a wireless bi-directionally transmitting communication network for multimedia applications characterized by the following properties:

- High bit-rate wireless transmission of media data via Media Points (MP) that connect mobile terminals to the Internet and to multimedia servers, which are locally distributed according to the expected user/terminal density, and in total do not enable a ubiquitous but only a discontinuous radio coverage and radio connectivity. An MP is realized by a base station with a broadband radio interface or by an FWR connected wirelessly to an AP.
- Continuous use of media (in spite of discontinuous radio coverage) whereby the application running on the MT in co-operation with the ISC of the WMS transfers data at each MP between the network and the MT as much as required for the expected processing and consumption within a planned time horizon, e.g., 20 min. By sending commands to the network, ISC will predict and determine what media data should be transmitted next to which MP to be predicted to be visited by the MT in the near future. In the meantime other local data is consumed or processed offline at the MT, and the results may be transferred to the network at the next MP.

An example scenario is shown in Fig. 2 with two MPs and their transmit and receive devices (transceivers) mounted on two gantries over the highway. Each MP is equipped with a control unit that is connected to the Internet or local multimedia server via the ISC of the network. The coverage area of an MP is depicted as a grey elliptic surface. As long as a vehicle resides in a coverage area, an MT in the vehicle can set up a connection with the local MP to request and download its application data wirelessly, or to send data, that has been prepared earlier, e.g., video recordings. When the vehicle leaves the coverage area or does not reside in any of the coverage areas, no radio link and thus no data transmission in absence of a contact to an MP exists anymore.

B. Provision of Personalized Data by Wireless Media System in City Areas

Fig. 3 depicts one envisioned scenario in a city area where a user/pedestrian can access the Internet via a WMS Access Point (AP) or Fixed Wireless Router (FWR) that is mounted invisibly at a street lamp. We consider such Media Point (MP) availability in the WMS city scenario, which is described in the following. There, many WMS

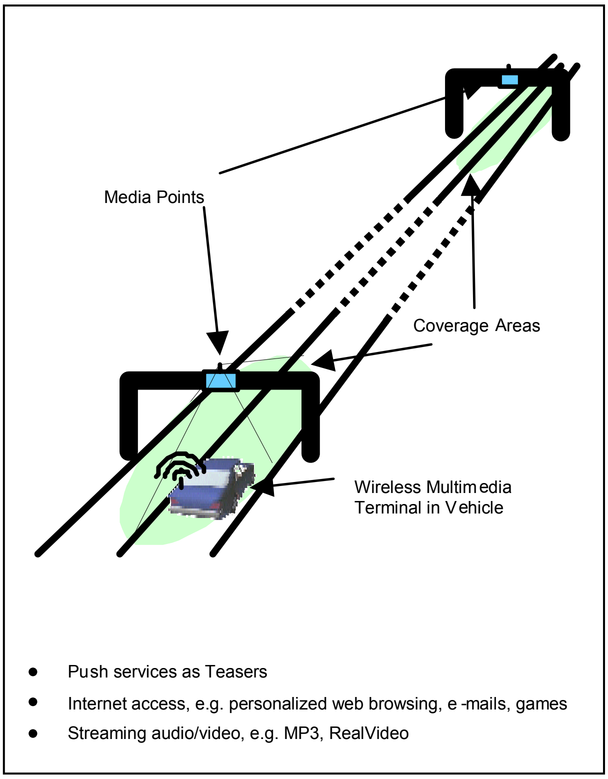

Figure 2. Continuous media use with discontinuous wireless data transmission for mobile terminals (highway scenario)

hotspots are available at strategic public areas. Users with suitable wireless equipments can “hop” from one MP island to another to get high-speed Internet access. One main issue WMS has to deal with is to be aware of the presence status and location of each MT. Personalized data like e-mails or subscribed multimedia contents (news, movie trailers, No.1 song of the week, etc.) should be brought or pushed to the user each time new data is available and the terminal is in reach of the WMS and the user is online (according to user preferences). Since the dwell time of the mobile user within an MP coverage area is limited and the push session could therefore be interrupted, the ISC of the WMS must support fast and reliable session re-establishment and resume at the next MP. To explain one of the technical approaches applied by the ISC we consider the WMS network architecture shown in Fig. 4.

The ISC acts as a local central entity, whose main functions are to monitor the presence status and location of each active user and to cache personalized user data. ISC is connected via broadband Intra- or Internet to a number of Media Point Controllers (MPCs) with large cache capacity. Each MPC administers a group of MPs that are usually located in close geographical proximity and establish the hotspots in city areas.

We consider the Session Initiation Protocol (SIP) [8] as a suitable protocol for handling the mobility management [9] of each user terminal and the signaling between the system components for push sessions. SIP has been standardized within the Internet Engineering Task Force (IETF) as a norm for the creation and termination of multimedia

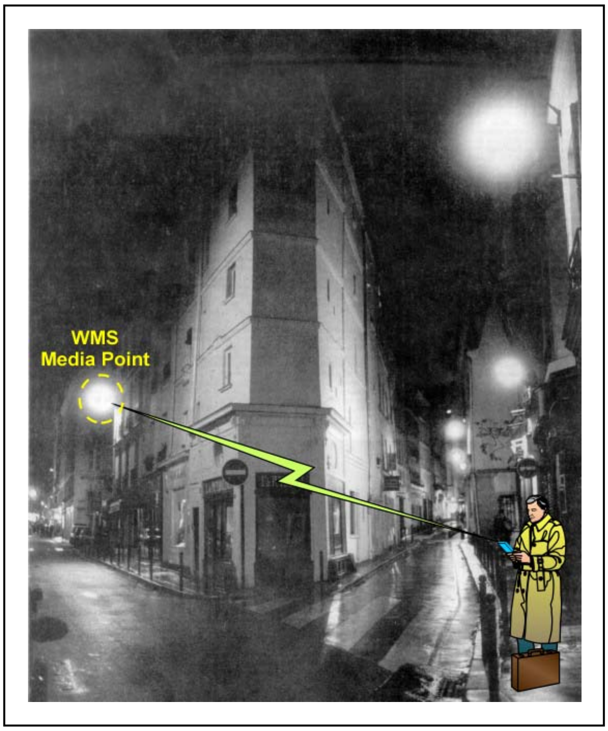

Figure 3. Media Points at hotspots in city areas

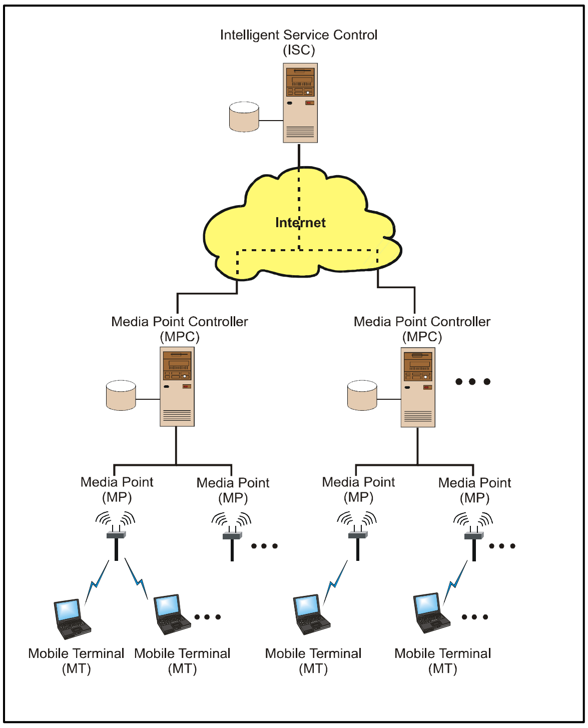

Figure 4. WMS network architecture

sessions. SIP is an application-layer control protocol designed to be independent of the lower-layer transport protocol.

Presently UDP is preferred as the default transport protocol but TCP is supported as well. Within the SIP context each (mobile) user is uniquely identified by his or her global SIP address. The SIP-based signaling in WMS is depicted in Fig. 5.

The presence status of each user terminal is stored in the so-called Presence Server (PS) [10]. The ISC requests for the user's presence status by sending a SUBSCRIBE message to PS on a regular basis as the message expires after a defined time. In case PS accepts the subscription, the current presence information [11] about the user is included in a NOTIFY message and sent back to the ISC. Whenever the presence status of the user changes the ISC will be notified. On the other side, after the mobile terminal (MT) has entered a hotspot area and set up a connection with an MP, the MT sends a REGISTER message to the administering MPC, which in turn forwards the message to PS. This way PS knows the current presence status of the user and via which MPC the user can be contacted.

If any personalized data for the online user is known available to the ISC, a new push session between the ISC and the MPC is initiated with an INVITE message. The session is described using the Session Description Protocol (SDP) [12]. After the session is established the ISC sends a list of data (by means of the MESSAGE method [13]) that the MPC can pull (download) and cache for the user. Instead of "true push" we apply the "smart pull" approach for WMS as it gives MPC more flexibility on how and when the data should be downloaded depending on the capability of each MPC, e.g., cache/storage size, computer power, link quality. The session is terminated by means of BYE message after all data has been pulled or after timeout. The identical SIP-controlled "smart pull" session is carried out between MPC and MT via the established wireless connection. Each time a data has been pulled by the MT an acknowledgment is sent back to the MPC and the ISC as well and the data is removed from their caches. In case that the "smart pull" session is interrupted, e.g., due to lost of AP connection, the MPC keeps the remaining data in its cache until timeout. The session can be resumed by the MT with the same MPC (e.g., via the same or another AP of the same group) or another MPC.

The strategy to cache user data in MPCs before it is pulled by the user becomes more advantageous when the link capacity between the ISC and MPCs is limited, e.g., due to high Internet traffic, and the data can be cached in an MPC before the user has entered one of the administered MP coverage zones. This requires

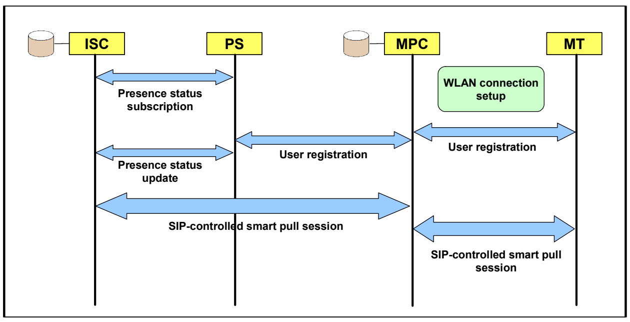

Figure 5. SIP-based signaling in WMS for pushing personalized data to a mobile terminal

sophisticated localization methods that can be realized by integrating the WMS with a mobile radio network.

C. *System Characteristics and Features*

In order to enable a continuous media use in spite of discontinuous radio coverage, the mobile terminals are equipped so that they can predict and determine in cooperation with the ISC in the network, which MPs they will pass by in a foreseeable time span next, and which media data they are going to download there. Since this prediction may only be made with a certain precision regarding the time and location, the network thus keeps the requested and other supplementary information available at several such MPs (i.e., their administering MPCs), that are foreseen by the terminals and the network as potential MPs to be passed by. Because the mobile terminals always maintain a virtual connection to the network (as required), they can refer to the preceding communication context at subsequent MPs/MPCs. The interworking between the wireless broadband terminals and the service control in network is coordinated so that the terminal's user gets the impression his or her terminal would always have a continuous connection to the network, although it is only connected via radio with the network from time to time. The mobile terminals possess a large internal storage capacity to receive or transfer the data wirelessly via an MP. If no free storage is available in the terminal to download data via an MP, the radio link will be released and communication will be resumed at one of the subsequent MPs by referring to the existing virtual connection.

Furthermore, by means of an extrapolation of the requested data or the corresponding application, e.g., playback of video clips or MP3 music, etc., the ISC extrapolates, which or what kind of data would be needed by the mobile terminal in the future, to assure a continuous data provision for the application. The data that has been requested by the terminal and made available by the network at a certain MP/MPC will be deleted if the terminal does not download the data within a certain period of time. The MPC can keep such media data, that are statistically very often requested by many mobile terminals, in its local storage permanently or at least for some time period, in order to minimize transmission delays between MPs and the Internet or multimedia servers via the ISC.

To support the prediction of the locations where supplementary data will have to be provided, the routes of the streets or highways are used to determine the reachable MPs, that are located on the streets. The terminals deliver terminal and vehicle specific data, such as GPS data, vehicle's speed or user-chosen route, etc., via the MPs to the network, which then evaluates the data and uses it in cooperation with the ISC for the prediction and determination of the MPs to be passed by the terminals/vehicles. For this purpose, information about the sequence of the MPs passed by the terminals so far, which is stored and updated regularly in the network, will also be evaluated in combination with the location specific characteristics of each MP. Further a regularly updated display on the mobile terminal may indicate the locations of the nearest MPs. The display on the terminal may inform the user or the application about the amount of application specific (personalized) data remaining on the storage and suggest a move direction for the mobile terminal, so that it could reach the next MP in time.

During a radio connection with an MP the terminal displays the actual capacity level of the local storage for the media data to the user, so that he or she could adjust the move speed to satisfy the data needs of the application. Other MPs near to the current MP will also be displayed on the terminal properly, so that the user, who can estimate his or her own move speed, can also estimate how much data he or she has to download at one time so that no lack of application data until arriving at the next MP occurs. Since not all MPs have the same transmission capacity and same traffic load (caused by the served terminals), the terminal's display informs also about the capacity and workload of the neighboring and other MPs, that are specifically requested by the terminal. This information might be used by the user or the terminal internal agent to organize the future route of the terminal and the consumption of the data stored locally.

To achieve this functionality, MP system protocols operate on top of the radio network and fixed telecommunication network (Internet) that allow a closely harmonized interaction between the network elements like mobile terminal, MP, ISC and Internet based servers.

IV. CONCLUSION

This paper proposes a service architecture for a wireless broadband system based on fixed infrastructure with multi-hop capabilities, see [14]. An Intelligent Service Control (ISC) mechanism is proposed, which hides the discontinuous radio coverage available from this Wireless Media System (WMS) from the users to a maximum degree so that the subscribers are provided with the contents they need within given time limits. It is shown that the Session Initiation Protocol (SIP) and its extensions can be applied to provide the session and mobility management of the WMS at the application level. Issues on integrating the 3G technologies into the WMS, e.g., regarding the terminal's localization or multi-access management, are subjects for further research.

ACKNOWLEDGMENT

The work presented in this paper has been supported by the German Federal Ministry of Education and Research in the Projects Multihop (01BU050), Multifunk (01BU901/3) and IPonAir (01BU163)[1].

REFERENCES

[1] B. Walke, "The Wireless Media System – a candidate next generation systems," Contribution to the 6th WWRF meeting, March 6-8, 2002, Tempe, Arizona, USA.

[2] B. Walke, "Spectrum issues for next generation cellular," Proc. Visions of the Wireless World, Workshop of the Wireless Strategic Initiative, December 12, 2000, Brussels, Belgium.

[1] The authors are responsible for the content of this publication.

[3] B. Walke, „HiperLAN/2 – Ein weitgehend in Deutschland entwickelter Standard," BMBF Symposium „Zukunftsperspektiven Mobilkommunikation", Juni 2000, Neu Ulm, Germany.

[4] B. Walke, "On the importance of WLANs for 3G cellular radio to become a success," Proc. Aachen Symposium on Signal Theory, Aachen University (RWTH), VDE Verlag Wiesbaden, Sept. 2001, available from http://www.comnets.rwth-aachen.de.

[5] B. Walke, M. P. Althoff, P. Seidenberg, "UMTS: A Comprehensive Introduction," John Wiley & Sons, Chichester, Weinheim, New York, Brisbane, 2002.

[6] B. H. Walke, V. Kumar, "Spectrum issues and new air interfaces," Computer Communications, August 2002, Elsevier Science Shannon, Ireland.

[7] B. Walke, "Mobile Radio Networks – Networking, Protocols and Traffic Performance," 2nd Edition, John Wiley & Sons, Chichester, Weinheim, New York, Brisbane, 2001.

[8] J. Rosenberg, et al, "SIP: Session Initiation Protocol," RFC 2543, February 2002.

[9] E. Wedlund, H. Schulzrinne, "Mobility Support using SIP," Second ACM/IEEE International Conference on Wireless and Mobile Multimedia (WoWMoM'99), Seattle, Washington, August, 1999.

[10] J. Rosenberg, et al, "SIP extensions for presence," Internet Draft, September 2001.

[11] H. Sugano, et al, "Common Presence and Instant Messaging (CPIM) Presence Information Data Format," Internet Draft, May 2002.

[12] M. Handley, V. Jacobson, "SDP: Session Description Protocol," RFC 2327, April 1998.

[13] B. Campbell (Editor), et al, "Session Initiation Protocol Extension for Instant Messaging," Internet Draft, June 2002.

[14] B. Walke, I. Herwono, R. Pabst, D. Schultz, "A mobile broadband system based on Fixed Wireless Routers," Proc. of the 7th WWRF workshop, Dec 2002, Eindhoven, Netherlands.