A MANAGEMENT ARCHITECTURE FOR ENABLING WIRELESS SYSTEM OPERATION IN THE B3G CONTEXT AND EXPLOITATION OF THE ABC CONCEPT*

Artemis Koutsorodi[1], Vera Stavroulaki[1], Panagiotis Demestichas[1, 2], Nikos Koutsouris[1], George Koundourakis[1], Dimitris Kouis[1], Nikolaos Mitrou[1]

[1]National Technical University of Athens, School of Electrical and Computer Engineering
9 Iroon Polytechniou St., Zographou 157 73 Athens, Greece
[2]University of Piraeus, Department of Technology Education and Digital Systems
185 34 Piraeus, Greece
e-mail: artemis@telecom.ntua.gr, vera@telecom.ntua.gr, pdemest@unipi.gr, nkoutsouris@telecom.ntua.gr, gkound@telecom.ntua.gr, kouis@telecom.ntua.gr, mitrou@softlab.ntua.gr

Abstract: **This paper presents a management architecture that enables wireless systems to operate in the Beyond 3G context, while allowing users to benefit from the Always Best Connected concept. It focuses on the structure and functionality of a general network and service management system for heterogeneous wireless access networks, capable of optimising service delivery and traffic load distribution in composite radio environments. Aspects concerning the communication between the user terminals and this management system receive special attention. A test case where the validity of this management architecture's functionality is verified, and consequent results, are also presented.**

1. Introduction

The evolution of wireless communication systems over the past years demonstrates a clear trend towards architectures that will support multiple radio access technologies and types of terminals, and that will incorporate functionality for the support of the Always Best Connected concept. The latter allows a person connectivity to applications using the devices and access technologies that best suit his needs, thereby combining the features of access technologies, such as WLAN, with cellular systems, such as UMTS, in order to provide an enhanced user experience for 3G and beyond. A challenging issue related to the above is the development of a management framework that will enable optimal system operation, thus allowing users to enjoy high-quality wireless services in a seamless way.

This paper is structured as follows: The second section defines and discusses the B3G and ABC concepts. The third section introduces a management architecture for enabling the realization of these concepts in a composite radio system, and outlines its main modules. The fourth and fifth sections elaborate on the functionality of the Session Manager and the Network Manager modules of the management system respectively. The sixth section presents a test case and results that serve to check the validity of this management architecture's functionality, and the final section concludes the paper.

2. Beyond 3G and Always Best Connected concepts

The next generation of mobile systems is expected to comprise heterogeneous networks consisting of diverse radio segments, able to host multimode wireless terminals, each of them capable of alternatively operating in each of the diverse radio segments available in the infrastructure. The different radio segments, or access technologies (e.g. WLANs, cellular and broadcast networks), will be interconnected by a backbone (e.g. an IP-based fixed network) and jointly operated in an optimized fashion that will allow for an improved overall resource management and network resource provisioning in reply to requests from the different service providers. These optimisation opportunities will increase the feasibility of providing advanced QoS-aware content services to the users of the multimode terminals over the composite network.

It is clear that the most salient feature of these evolving systems will be the multiplicity of access technologies and terminals, that will allow users to enjoy wireless services at any time, at any place. This constitutes the *always best connected* (ABC) concept: being connected at any point in time to the best available access network. The latter is defined as the network that best suits the user's needs at that particular moment and may be a function of the user's personal preferences, of the different operator's policies and business agreements and of the infrastructure's available network resources. The application of the ABC concept to an environment of multiple access technologies allows for the beneficial combination of the coverage of cellular and broadcast

* This work is partially funded by the Commission of the European Communities, under the Fifth Framework Program, within the IST project CREDO (IST-2001-33093: Composite Radio and Enhanced Service Delivery for the Olympics).

systems with the flexibility and high bandwidth of WLAN systems.

In this context, the exploitation of the composite radio infrastructure requires an innovative Network and Service Management System (NSMS), capable of driving the user terminals to the appropriate quality level and network for each service, in a transparent manner. Figure 1 depicts a general composite radio architecture that includes such a management component.

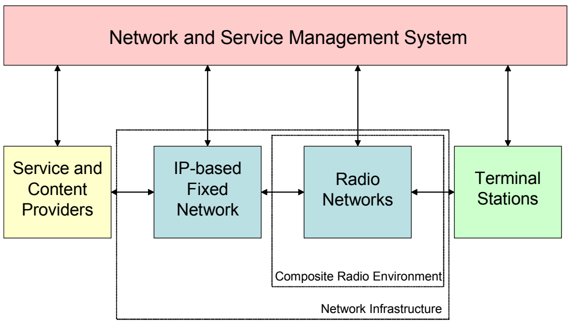

Figure 1. Main components in a composite radio architecture.

3. Management architecture for enabling the B3G and ABC concepts

Optimal system operation is to be achieved through the joint contributions of both the network and the terminals: On the network's side, the NSMS will perform tasks such as the balancing of the traffic load between the different radio segments, and the general coordination thereof – the goal being optimal joint resource management. In more detail, the NSMS should be capable of monitoring and analysing the performance and the QoS levels provided by the different network elements, of dynamically calculating and imposing the optimal traffic load distribution, and of inter-working with the user terminals. Figure 2 depicts the main modules of the NSMS entity, that carry out the above-mentioned tasks, namely the Session Manager and the Network Manager.

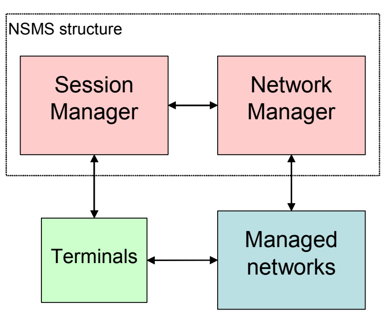

Figure 2. Main modules of the network management component (NSMS).

The Session Manager addresses a short-term optimisation problem, targeted to the assignment of the user terminal to a specific network, thus enabling the selection of the appropriate radio technology through which services can be obtained efficiently in terms of cost and QoS, in near real time. The Network Manager, on the other hand, addresses a mid-term optimisation problem, targeted to the splitting of the aggregate traffic load to the underlying networks, and to the assignment of traffic to QoS levels.

On the terminal's side, a Terminal Station Management System (TSMS) that will interact with the NSMS is required. The purpose of that interaction will be making the optimal selection of the appropriate radio segment to which the terminal will eventually be assigned. This interaction will also ensure that the network's and the terminal's estimate of radio conditions and QoS levels in the system are beneficially combined for making an educated selection of the appropriate radio technologies through which services can be obtained as efficiently as possible. Thus, both the network and the terminals contribute intelligence towards optimal system operation.

The interaction between the TSMS and the NSMS is not expected to be limited to the initial radio selection, but will rather be maintained throughout the entire stay of the terminal in the composite radio environment. The purpose of this communication will be the handling of events originating either from the terminal, such as requests for the provision of additional services, information regarding the radio conditions in the area or the QoS levels associated with the services currently running on that terminal, or from the network, such as the re-distribution of terminals across radio segments for the purpose of balancing the overall load. Both kinds of events may result in terminals being subject to a vertical handover.

Another noteworthy feature of this approach is the support of advanced QoS-aware services. It is assumed that the users of the composite network will have access to advanced IP-based services, obtained through application client software running on top of the wireless terminals and accessing content from application servers over the network. These services will be offered at multiple QoS levels. The accessibility to a service and its particular QoS level will be determined by the user's profile.

The following sections discuss the functionality of the Session Manager and the Network Manager of the NSMS in more detail.

4. The Session Manager

4.1. Functionality for supporting the ABC concept

In an environment of multiple access technologies, where the user is connected through the best available access technology and is able to seamlessly move between these network technologies while maintaining connections to application servers, functions are often considered to be broken into distinct blocks [1]. These include access discovery, access selection, authentication/authorization/accounting support and mobility management among others, as illustrated in Figure 3, based on an approach presented in [1].

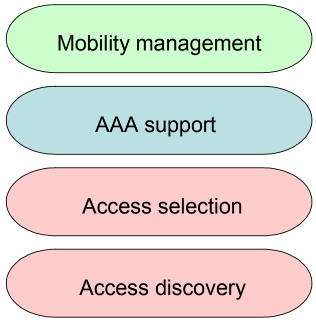

Figure 3. Technical components in an environment of multiple access technologies, as presented in [1].

Access discovery involves finding the available access networks and periodically performing new access discovery to find out if a better alternative has become available. An important issue involving these tasks is the collection of statistics of the different access networks.

Access selection, as its name implies, involves the process of deciding over which access technology the terminal should connect. In the case of terminal-based access selection the terminal needs to be informed of the capabilities of the access networks. In the case of network-based support for access selection, a network management system is required in order to provide network-specific information and to achieve load-balancing and to maximize total system throughput.

Authentication, authorization and accounting support involves the tasks of verifying the user's identity, establishing which services the user is entitled to and collecting all the necessary information to bill the user for these services. In a broader sense, the adopted AAA solution would also be responsible for the general handling of the user's profile. The latter is set up when a user subscribes to a service and contains information such as the user's personal preferences, his or her credentials for authentication purposes and accounting data.

Mobility management refers to the ability to maintain a session when the terminal moves between different access technologies and to reach the user on his or her current network and terminal at any point in time. Mobility management may also be enhanced to include the ability to maintain a session when the user moves between different devices, in which case appropriate content adaptation, needed in order to adapt to different device capabilities, becomes a key issue.

The development of protocol solutions for access discovery and access selection is a topic of ongoing research (e.g. [2]), as no specific standard has yet been developed to cover these two functions in a systematic way. The Session Manager component of the management architecture discussed in this paper is responsible for the execution of these two functions and fills the above-mentioned gap by adopting a simple signalling protocol, used between the network management entity and the TSMS, a terminal middleware. This solution does not preclude the potential merging of interfaces, in the future, in a common protocol that will cover all aspects of the terminal – network communication.

As for authentication, authorization and accounting solutions, the Internet Engineering Task Force is currently developing the standard for the next-generation AAA protocol, which is called Diameter. This protocol provides a number of enhancements over the existing protocol for remote authentication dial-in user service (RADIUS).

Where mobility management is concerned, session continuity, which refers to the ability to maintain a session when the terminal moves between different access technologies, may be provided for through the use of Mobile IP, an IP layer solution aimed at making movements on the IP layer transparent to higher protocol layers. User reachability, which refers to the ability to reach the user on his or her current network and terminal at any point in time, may be provided for through the use of the Session Initiation Protocol (SIP). More specifically, SIP is an application layer control protocol, whose main functionalities are user location, user availability, user capabilities, application level session set-up and application level session management [2].

4.2. TSMS – Session Manager information flow

The TSMS – Session Manager communication is of paramount importance for the overall functionality of the composite radio system. It enables the realization of the ABC concept by providing the information that is essential for the execution of the access discovery and access selection functions discussed in the previous subsection. The present subsection gives an overview of the TSMS – Session Manager information flow.

The protocol through which the TSMS and the Session Manager will communicate must perform the following functions:

- Upon initialization, the TSMS shall use it to request information regarding the services that the user of the terminal has subscribed to. The Session Manager shall in turn use the protocol to issue a reply to this request.
- The TSMS shall use it to request a service. The Session Manager shall in turn use the protocol to accept or to deny this request, and to specify the network to which the TSMS should switch for the provision of this service, in case the corresponding request has been accepted. This function constitutes one of the most basic aspects of the communication between the TSMS and the Session Manager.
- The TSMS shall use the protocol to notify the Session Manager that a service has stopped running on a terminal, and the Session Manager may use it to explicitly ask the TSMS to stop a service.
- The TSMS shall use the protocol to periodically report to the Session Manager the list of available networks in the terminal and the list of services that the terminal is running. The TSMS shall also

report to the Session Manager a set of status parameters of the terminal and of each of the available networks in the infrastructure (for example bit rate, received power level, etc.). The protocol should support the transmission of these quality reports both on the TSMS's own initiative (for instance when the perceived service quality is reduced) and upon a request from the Session Manager.

- The TSMS shall use the protocol to issue keep-alive probes, so that the Session Manager can be aware of an unexpected termination of the terminal, that would otherwise go unnoticed. This information is critical to the Session Manager since the validity of its calculations relies heavily on the accuracy of the information it receives regarding the number of terminals that are present in the network and the services that are associated with these terminals.

- The Session Manager shall use the protocol to instruct the TSMS to perform a vertical handover to a different network. In emergency situations, such as overall network degradation, the protocol may also be used to request quality parameters from the terminal. In such a case, the TSMS shall in turn use the protocol to provide the Session Manager with the requested measurements reflecting the quality level at which each service is provided to the particular terminal.

4.3. Session Manager core operation

The Session Manager is the NSMS module responsible for performing all operations concerning the communication between the NSMS and the terminal. It composes, sends, receives and processes the messages through which this communication is achieved, according to the adopted communication protocol. The Session Manager holds information about the active terminals that are served by each network, and also about the quality level assigned to them. Based on that information and on consequent calculations, it issues recommendations to the terminals on choosing the best available network for the provision of a particular service. Thus, the Session Manager addresses a short-term optimisation problem, targeted to the assignment of the user terminal to a specific network. The solution of this optimisation problem enables the sophisticated selection of the appropriate radio technology, for a specific user, through which services can be obtained efficiently in terms of cost and QoS, in near real time.

The optimisation problem addressed by the Session Manager relies on the following input data: (a) the set of services the user is requesting and the corresponding set of quality levels at which these services are requested; (b) the profile of the user requesting the set of services. This includes parameters such as the maximum price that the user is willing to pay for the requested services; (c) the network policies. This mainly involves the cost deriving from the assignment of the user demand to the several quality levels.

The optimisation process carried out by the Session Manager should result in an allocation of the requested services to specific quality levels, and in an allocation of the requested services to specific networks. The calculation of these allocations should optimise an objective function, which is associated with the quality levels at which each service will be provided, and the utility deriving from the assignment of the user demand to high quality levels. These allocations are bound to certain constraints, such as the capabilities of the user terminal, or the limit to the overall price that the user is willing to pay during usage of the composite radio system. The Session Manager optimization process is illustrated in Figure 4.

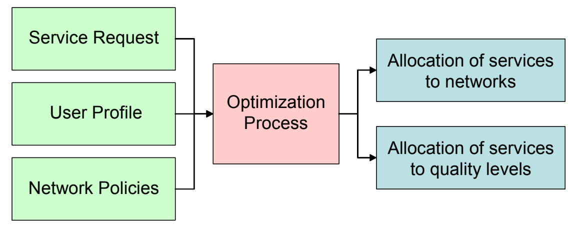

Figure 4. Session Manager optimization process

4.4. Software aspects

As previously stated, the TSMS – NSMS communication constitutes a cardinal part of the overall system management process. An interesting issue involves the software aspects regarding the aforementioned communication. This subsection presents a first approach regarding this issue.

The selected protocol for the communication between the TSMS and the Session Manager of the NSMS must be able to carry out all the functions outlined in subsection B in a simple way. This means that it should perform the tasks of requesting, allowing, denying and stopping a service, the task of requesting and issuing information regarding the terminal's quality parameters or the services a user is subscribed to and, finally, the task of instructing a terminal to perform vertical handover using as few messages as possible and avoiding redundancy in the exchanged information. It also means that the protocol should be able to carry all the necessary types of data structures, encode IP addresses, network lists and network types and, ideally, introduce as little overhead as possible. It can be viewed as a signalling protocol, since data transported in its messages is not application data. Thus, it should use as few resources as possible, since resources should be available for the user.

Another important requirement involves the speed with which this signalling information is relayed across the network: some of the messages that need to be exchanged can be urgent, like the notification of network unavailability, or the request of a network switch. It is important to note that typical round trip times in GPRS connections can be of about half a second, so the number of required packet exchanges in order to send a message through the TSMS – Session Manager protocol must be as small as possible.

The selected protocol should also provide a mechanism for the authentication of the exchanged messages. Authentication is used to prove the identity of the sender and to guarantee that the message has not been modified in transit. It should also support anti-replay protection. This is important for two reasons: First, from the security point of view, an eavesdropper could intercept a message and send it again without modifying it, thus causing confusion in the system. Second, it is also possible for the network to duplicate packets because of retransmissions. The protocol must also support acknowledgements for some messages, to guarantee that they arrive. Note, however, that not all types of messages may require acknowledgements. Finally, it must provide a solution for communication when one of the communicating entities is behind a NAT gateway.

Another important point is the fact that the TSMS – Session Manager interaction can easily be viewed as a client - server model, with the TSMS acting as the client and issuing requests and reports, and the Session Manager acting as the server, issuing replies and notifications. This observation, combined with the above-mentioned requirements, could lead to the decision to define this protocol over UDP, as presented in [4]. UDP offers a way of sending encapsulated IP datagrams without having to establish a connection between the parties exchanging the information. In general, UDP is characterized by speed and simplicity, its most significant downside being the lack of reliability where retransmission and ordering of packets are concerned. TCP, on the other hand, effectively deals with both of these issues at the cost of speed. In addition, many client - server applications, which involve requests from the client and subsequent replies from the server, prefer using UDP, rather than another transport layer protocol, for example TCP, that would have to go through the trouble of establishing and releasing a connection. It should be noted however that the adopted protocol for the TSMS – Session Manager interaction could just as well be defined over TCP, or over a higher-level protocol. In any case, the underlying philosophy should be the same.

As mentioned in subsection B, the TSMS – Session Manager interface is also used to report to the latter a set of status parameters of the terminal and of each network. The status parameters that are network-specific refer to the physical layer (for instance the received power level) and the data link layer (for instance bit rate) of the different radio segments, as well as to the network / IP layer (for instance upstream / downstream IP layer bit rate). In this sense, this interface offers a means for performing a network monitoring task, much like the SNMP protocol does: SNMP offers a systematic way of monitoring and managing a network. Managed nodes in the network must be able to locally run a managing process, an SNMP agent, that will keep a local data base of variables describing its status and history. The total of all variables (objects) in the network is given in the Management Information Base (MIB). The managing station interacts with SNMP agents through the SNMP protocol, which allows the managing station to ask about the status of an agent's local objects and to change them if necessary.

The SNMP model assumes that every managed node is able to internally run an SNMP agent. This however may not be true of some devices. In general, although SNMP offers a wide range of network monitoring and managing capabilities, providing a large amount of information from all layers, it is a bulky protocol that is impractical to implement in the device scale of laptops and portable phones.

5. The Network Manager

The Network Manager is responsible for the monitoring of the managed network infrastructure. More specifically it monitors and analyses the statistical performance and QoS levels provided by the network elements (segments) of the managed infrastructure, and the associated requirements originating from the service area (environment conditions, e.g., traffic load, mobility levels, etc.). It also assesses the relevant network and service-level performance, and dynamically finds and imposes the appropriate traffic distribution, through which the service management requests, and service area conditions are handled in the most cost-efficient manner. The Network Manager consists of the following entities: Resource Brokerage, Service Management, and Monitoring and Configuration (Figure 5) ([5],[6]).

The general functionality of the Resource Brokerage sub-component is to co-ordinate all the other entities of the NSMS in order to handle various conditions, such as congestion in a certain service area. The Service Management entity provides optimisation functionality for determining the appropriate service configuration (allocation to QoS levels) and aggregate traffic distribution (allocation to networks). The operation of these two sub-components is independent from the underlying radio access technology. The Monitoring and Configuration entity on the other hand, depends on the radio access technology. This latter component provides auxiliary functionality for handling new service area conditions or management requests. It mainly monitors the underlying network segments and forwards the related information to the Resource Brokerage entity.

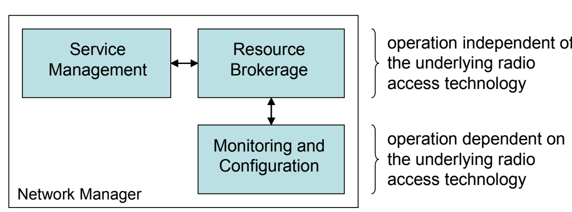

Figure 5. Network Manager sub-components

6. Results

A laboratory environment involving real network segments and innovative terminals was used for conducting a series of experiments validating the Session Manager functionality and its interaction with the terminal as well as the operation of the NSMS in general. The laboratory composite network infrastructure, comprises GPRS, Wireless Local Area Network (WLAN) and Digital Video Broadcasting – Terrestrial (DVB-T). The GPRS network consists of a GPRS cell with 8 time slots (TS) available for GPRS usage with CS-2 modulation type. A WLAN IEEE 802.11b Access Point also covers the area. The DVB-T network consists of a DVB-T gateway, in charge of multiplexing the packet traffic from an Ethernet interface into the DVB-T signal, a DVB-T modulator and a RF converter. The output rate for IP services was set at 10Mbps through proper configuration of the DVB-T gateway. A MIPv4 platform with a GPRS NAT handling module offered to the terminals seamless handover support between the various network segments. Three types of services have been assumed, namely Video Streaming, Sports Event and Generic Internet Service Provision (e.g. Web browsing). Figure 6 provides the two reference quality levels per service.

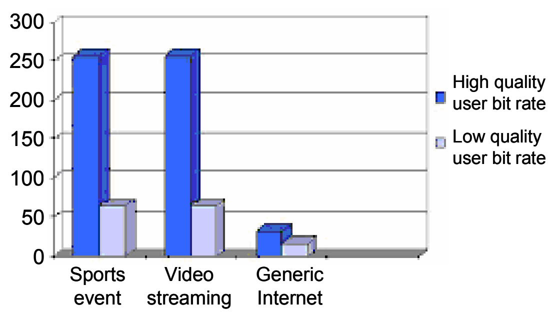

Figure 6. Quality of Service reference levels

In the results presented here, the initiation of the test case scenarios is made through the GPRS network. In other words, the GPRS network is considered the "Home" network. However, the results obtained can be considered generic with respect to which underlying wireless access technology is considered the "Home" network. Figure 7 depicts a snapshot of part of the graphical user interface of the TSMS contacting the Session Manager in order to request a Generic Internet Service. The Session Manager indicates the appropriate network (GPRS) and quality level (high) which results in the TSMS remaining at the GPRS network and starting the service.

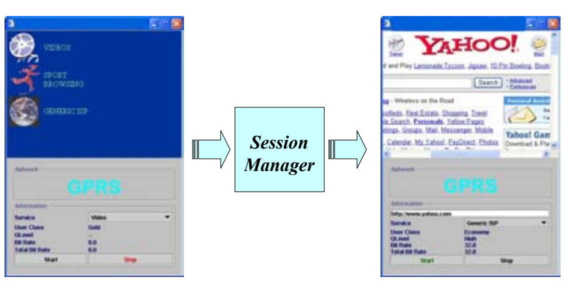

Figure 7. Result of Generic Internet Service Request

In the sequel (Figure 8), the TSMS requests a more demanding service, i.e. Video Streaming service. The Session Manager addresses the short-term optimisation problem discussed earlier and instructs the terminal to switch to the WLAN network in order to obtain both services at the highest possible quality level.

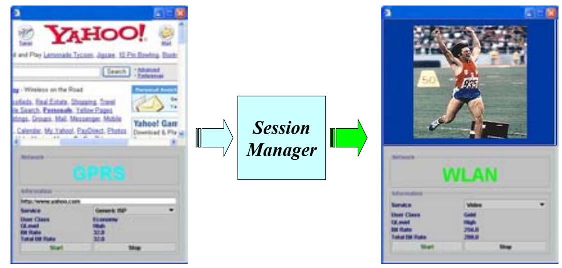

Figure 8. Result of additional request for Video Streaming service

The results depicted above are obtained when the status of the GPRS and WLAN networks is such that allows the efficient, in terms of QoS, provision of services. Alternatively, a situation may occur where for example, although the terminal requests a low demanding service (e.g. Web browsing), the GPRS Network is not available. This will result in the Session Manager addressing the short-term optimisation problem, taking into account that the available networks are now only WLAN and DVB. It should be noted that for the test case discussed here, the Session Manager, through network cost parameter set up, was configured to consider the WLAN network as the lower cost network (compared to DVB). Therefore, in a case where GPRS is unavailable, as the one described above, and the user requests Generic Internet, the terminal is instructed to switch to the WLAN network, as illustrated in Figure 9.

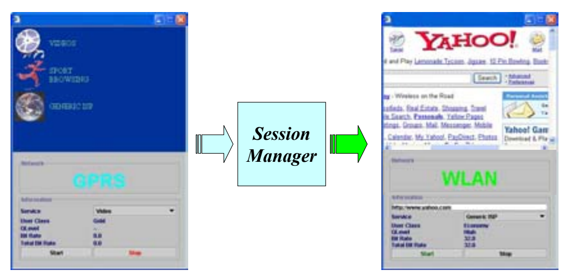

Figure 9. Result of Generic Internet Service Request

The results introduced in this subsection are indicative of the functionality of the Session Manager and its interaction with the terminal.

7. Conclusions

This paper presented a management architecture that enables wireless systems to operate in the Beyond 3G context, while allowing users to benefit from the Always Best Connected concept. It described the structure and functionality of a network and service management system (NSMS) for heterogeneous wireless access networks, capable of optimising service delivery and traffic load distribution in composite radio environments. It elaborated on the core functionality of each of the NSMS's main modules, namely the Session Manager and the Network Manager, and gave special focus on issues regarding the communication between the user terminal and the Session Manager. A test case where the validity of this management architecture's, and especially the Session Manager's, functionality is verified, and consequent results, were presented.

REFERENCES

[1] E. Gustafsson and A. Jonsson, "Always Best Connected", IEEE Wireless Communications, February 2003.

[2] E. Njedjou, P. Bertin, P. Reynolds, "Motivation for Network Controlled Handoffs using IP Mobility between heterogeneous Wireless Access Networks", IETF Internet Draft, draft-njedjou-inter-an-handoffs-00.txt, June 2003, work in progress.

[3] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[4] M. Catalina Gallego, P. Roux, "Candidate Network Selection in Composite Radio Environments", accepted for publication in HET-NETs '03 Technical Proceedings, July 2003.

[5] G. Koundourakis, N. Koutsouris, V. Stavroulaki, L. Papadopoulou, V. Tountopoulos, D. Kouis, P. Demestichas, N. Mitrou, "Network and Service Management System for Optimising Service Delivery and Traffic Distribution in Composite Radio Environments" In Proc. IST Mobile &Wireless Telecommunications Summit 2003, Aveiro, Portugal, June 2003.

[6] P. Demestichas, N. Koutsouris, G. Koundourakis, K. Tsagkaris, A. Oikonomou, V. Stavroulaki, L. Papadopoulou, M. Theologou, G. Vivier, K. El Khazen, "Management of Networks and Services in a Composite Radio Context", IEEE Wireless Communications Magazine, August 2003 (Evolution towards 4G Mobile Communication Systems).