(R)Evolution Toward 4G Mobile Communication Systems

MANAGEMENT OF NETWORKS AND SERVICES IN A COMPOSITE RADIO CONTEXT

PANAGIOTIS DEMESTICHAS,
NATIONAL TECHNICAL UNIVERSITY OF ATHENS AND UNIVERSITY OF PIRAEUS
NIKOLAOS KOUTSOURIS, GEORGE KOUNDOURAKIS, Konstantinos Tsagkaris, Antonis Oikonomou, Vera Stavroulaki, Louisa Papadopoulou, and Michael E. Theologou,
NATIONAL TECHNICAL UNIVERSITY OF ATHENS
GUILLAUME VIVIER AND KARIM EL KHazen, MOTOROLA LABS, PARIS

ABSTRACT

This article adopts the assumption that different wireless access technologies (e.g., UMTS, HIPERLAN2, and DVB-T) can be cooperating components of a composite radio infrastructure. The coordinated management of services and network in this heterogeneous context is a complex task. This article presents a service and network resource management platform for wireless systems that operate in a composite radio context. Aspects addressed are the requirements, high-level platform design, functionality of the components that conduct monitoring, resource brokerage, service management, and network resource optimization. Indicative results are presented and concluding remarks provided.

BACKGROUND

Wireless communications continues to attract immense research and development effort. From a technology perspective, the main evolutions are the enhancement of mobile infrastructures toward the second generation plus (2.5G) and third-generation (3G) era [1, 2], the introduction of broadband radio access networks (BRANs) [3], and the advent of digital video broadcasting (DVB) [4]. Moreover, a more recent trend often called wireless beyond 3G or 4G wireless communications, is to assume that mobile, BRAN, and DVB systems can cooperate in a composite radio (CR) infrastructure [5].

Wireless systems should address the following in order to be successful:

• Cost-effective provision of new services at the appropriate quality of service (QoS) levels
• Efficient handling of versatile service area conditions (e.g., hotspots caused by time-variant traffic loads and mobility)

According to the CR concept, a network provider (NP) can rely on diverse radio technologies to efficiently cover subsets of the service area (i.e., service area regions). This may mean that the NP either possesses licenses for deploying and operating diverse radio systems (tight integration between radio technologies) or cooperates with other NPs that own alternate radio networks (loose integration between radio technologies).

Efficient coverage means offering the highest possible QoS levels at adequate capacity volumes in a cost-effective manner. Therefore, an NP can choose, in a certain service area region and time of day, instead of rejecting users or degrading their QoS levels, to direct them to an alternate radio technology (which may belong to an affiliated NP). In this respect, users can obtain their services through different wireless technologies, based on the assumption that the supporting radio technology can be irrelevant to them and to service providers (SPs), as long as cost and QoS criteria are met. Multimode terminals, through the software/reconfigurable radio concepts [6], can enable the CR concept.

The exploitation of a wireless system, operating in a CR context, requires upgraded service and network resource management (SNRM) capabilities. In this direction, the aim of this article is an SNRM platform [5] in a CR context that comprises Universal Mobile Telecommunications System (UMTS), High Performance LAN 2 (HIPERLAN2) [7], and terrestrial DVB (DVB-T) wireless access systems and IP-based fixed networks.

The discussion evolves as follows. We provide the requirements for the SNRM platform. We introduce a generic SNRM architecture, and describe its operation in a CR context. We pre-
sent the design of the main components in the SNRM architecture. We then present results and concluding remarks, respectively.

**Requirements and High-Level Design**

An SNRM platform, for a wireless system that can operate in a CR context, should be capable of three main aspects:

- **Handling new service area conditions (e.g., new traffic demand patterns).** This requires efficient monitoring of the wireless system (component of the CR infrastructure). It constitutes a reactive mode of operation.
- **Handling service management related requests.** This can be interpreted as the determination of the appropriate QoS levels and networks (of the CR infrastructure) for the provision of (potentially new) services, within certain service area regions and time zones. It constitutes a proactive mode of operation.
- **Handling resource brokerage requests.** Coordinating in the handling of new service area conditions, or service management related requests, originating in other networks of the CR infrastructure.

**High-Level Design and Operation Architecture**

A generic SNRM architecture, targeted to the requirements above, consists of three entities:

- **Monitoring, service management interworking, and resource brokerage (MSRB).** It provides auxiliary functionality for initiating and supporting the handling of new service area conditions or service management requests. Resource brokerage is a capability imposed by the CR concept.
- **Resource management strategies (RMS).** It provides the optimization functionality for finding the appropriate system reconfigurations, coordinated with the CR infrastructure, for handling new service area conditions or service management requests.
- **Network and environment simulator (NES).** It enables validation of some management decisions prior to their application to the network, offline, testing and demonstration.

Figure 1a presents a pattern for deploying this generic SNRM architecture in a CR context. It depicts three independent networks (wireless access technologies) belonging to three different operators. The fixed networks are assumed to be IP-based. The SNRM platform in each network manages a specific radio technology. However, the platforms can cooperate. The existence of different platforms is a realistic approach, enabling NPs to maintain private information (e.g., the exact network structure).

**Operation**

Figure 1b depicts a sample scenario according to which the entities of three SNRM platforms collaborate. The NES components are omitted for simplicity. Our discussion is generic with respect to the timescale of the collaboration, which can be an automated dynamic process or a more long-term business-level process comprising offline phases.

In the scenario, the UMTS network has to handle a new service area condition or a service management request. The network that faces such a trigger (event) is called originating (in this case, the UMTS network), the other networks cooperating (in this case, HIPERLAN2 and DVB-T networks).

Based on the above, the general aims of the platform's operation are the following:

- To evaluate the demand volume that should be accommodated to adequately handle the event
- To find the QoS levels that can be offered
- To select and configure the networks that can support the demand at the selected QoS levels

The sample scenario evolves as follows:

- **The MSRB-U identifies a new service area condition or receives a service management request.** The demand volume that should be accommodated is evaluated, and the service area regions that will be affected are found.
- **The MSRB-U acquires the status of the network and of the service level agreements (SLAs) in the affected service area regions.**
- **The cooperating networks are asked whether they can participate in the handling of the event.** Essentially, the SNRM entities of each network (MSRB-H/RMS-H, and MSRB-D/RMS-D) prepare offers for describing the terms of the participation.
- **The RMS-U is invoked.** It conducts the service configuration and distribution phase, in which it consolidates the best assignments of the demand volume to QoS levels and networks. The derived solution is returned to the MSRB-U entity.
- **Ideally, the solution reached above is found acceptable by the MSRB-U.**
- **The networks of the CR infrastructure are configured in accordance with the solution accepted in the previous phase.**

**Trigger Identification (Service Management Request — New Service Area Condition)**

This component initiates the SNRM operation. The service management request part is an application programming interface (API) offered to legacy management/planning software, which is also accessible to a human operator through a Web-based interface. The new service area condition part exploits standard information extracted from networks (explained below), so as not to impact normal control information exchange.

The trigger description provides fundamental information for proper handling of the service management request or new service area condition. Such information is the following:

- **The involved service (or set of services) it should be noted that the MSRB-U, MSRB-H and MSRB-D target only one network (technology), and therefore, each operator has limited access to the details of other networks. In part the component functionality is independent from the underlying technology.**
Each permissible QoS level is associated with a set of parameters, and for each parameter there is a target value, or range of values. For fairness reasons, the users of each class should receive (as much as possible) the same QoS levels, in the involved service area regions and time zone.

The user classes to which each service is offered
- For each user class, the estimated volume of users and profile/policy information
- The involved service area regions and time-related information

The profile/policy information specifies:
- The permissible QoS levels and candidate networks per service
- The importance (utility) [8] and the allowable cost associated with each QoS level of each service
- The traffic and mobility behavior, per service, of a typical user of the class

Each permissible QoS level is associated with a set of parameters, and for each parameter there is a target value or range of values. For fairness reasons, the users of each class should receive (as much as possible) the same QoS levels in the involved service area regions and time zone. The terminal capabilities (and involved services) provide the candidate network technologies by which a user class can be served.

The following information is in response to a service management request or new service area condition:
- The QoS levels that can be offered per user class (subset of the set of permissible ones)
• The distribution of the demand volume to networks
• The costs of the assignments to QoS levels and networks
The information above will be valid in the involved service area regions and time zones.

**Resource Brokerage**
This capability is imposed by the CR concept. It enables cooperation between the NPs of the CR infrastructure [9]. Specifically, the NPs of the CR infrastructure should be capable of exchanging and negotiating with a set of offers.

An offer request should indicate the services, QoS levels, demand volume, involved service area regions, and time zones. In response, each offer should specify the QoS levels supported per service and the cost of the resources. In this respect, a flexible model, enabling discounts or congestion-dependent charging, is to consider each offer as a set of triples \((b_{\text{min}}, b_{\text{max}}, p_{\text{bu}})\). If the bandwidth reserved is \(b_r\), where \(b_{\text{min}} \leq b_r \leq b_{\text{max}}\), the corresponding cost, imposed by the network making the offer, will be \(b_r \cdot p_{\text{bu}}\), where \(p_{\text{bu}}\) is a price imposed per bandwidth unit reserved.

A second capability of this component is resource reservation, which takes place after the distribution of the demand to the CR infrastructure.

**Status Acquisition — Network Monitoring**
This component monitors the status of the established SLAs and that of the managed network (component of the CR infrastructure).

The SLA status description provides information on the basis of service and user class. The network status description aggregates information. It reflects the current configuration, load, and performance of distinct elements or segments of the network, covering certain service area regions.

This is the MSRB component that depends on the underlying network technology. The component relies on the collection of information (through interactions with legacy network/element management systems) from the CR infrastructure. The rest of this subsection presents some requirements for each of the underlying networks.

**UMTS.** As the air interface can be the main bottleneck, of primary importance is the information from cells (Node Bs). The following is typical cell-level information (assuming that cells can be equipped with more than one 5 MHz carrier):

- The demand volume and QoS offered (e.g., bit rate), per service and user class, on the uplink and downlink, on each carrier. Such information is the basis for extracting higher-level information (e.g., SLA status).
- The loading factors (interference patterns) on the uplink and downlink, per carrier. This is a compact expression of the aggregate cell load, useful in capacity calculations.

**HIPERLAN2.** Information from access points (APs) is collected. Important is the extraction of the demand volume and QoS per service and user class, as in the UMTS case. Equally important is the characterization of each AP’s radio link status on the uplink and downlink. This is achieved through the, uplink and downlink carrier to interference ratio (C/I) sensed on the frequency used, bit rate supported, and overall link utilization [7].

**DVB.** DVB-T networks are primarily designed for the delivery of audio/video/data content related to TV services. Standardization allows the encapsulation of (TV service unrelated) IP data in the transport stream [4]. The focus is on collecting load and QoS statistics, per IP service.
The RMS entity conducts two general tasks: service configuration and distribution (SCD) to the CR infrastructure; and network configuration (NC) for accommodating the demand and the QoS levels that have derived from the SCD.

**RMS Design**

The RMS entity conducts two general tasks: service configuration and distribution (SCD) to the CR infrastructure [10]; and network configuration (NC) [11,12] to accommodate the demand and QoS levels derived from the SCD.

**Service Configuration and Distribution**

This optimization problem enables the originating NP to address the following objectives:
- To allocate the user classes to QoS levels (service configuration task)
- To allocate the demand volume to networks of the CR infrastructure (service distribution task)

Figure 3a is the high-level problem description. The input consists of the trigger description, the status of the originating network, and the offers of the cooperating networks.

The SCD solution should optimize an objective function that takes into account the utility volume deriving from the assignment of the demand volume to QoS levels, and the cost deriving from the assignment of the demand volume to QoS levels and networks.

The utility volume differentiates the importance of QoS levels. It drives the originating NP to try to allocate user classes to the most “preferred” QoS levels (from the set of permissible ones). The cost of the assignments is associated with the following aspects:
- The cost of serving demand portions at the selected QoS levels in the originating network
- The cost that the selected cooperating NPs will impose to the originating network for support of their assigned demand portions at the selected QoS levels

The constraints of the SCD problem fall into two general categories. The first includes those that derive from the profiles/policies. User classes (and services) should be assigned to QoS levels and networks that are compliant with their profiles. The second category includes those deriving from the status and offers of the networks of the CR infrastructure.

**Network Configuration**

The objective of this component is to find the best network configurations to accommodate the demand and QoS assignment proposed by the SCD. The proposed configurations are applied to the elements that serve or are near the affected service area regions and are valid for specific time zones.

**UMTS**

Essentially, the component configures the admission control and load control operations [1]. The component conducts two main tasks (Fig. 3b). First, a system-level decision is made regarding the distribution of the demand (assigned by the SCD) to the 5 MHz carriers. The objective of this phase is to bring the system level performance close to given operating points.
DVB networks cover 100 percent of the service area. HIPERLAN2 varies between service area regions. Each UMTS cell is equipped with two carriers. The target is to operate UMTS cells at no more than 75 percent loading on the downlink and 55 percent on the uplink, per carrier. Therefore, it can be considered that each cell can provide up to 110 voice connections, which means that up to 90 Erlangs of equivalent-to-voice traffic load can be supported with 1 percent blocking probability.

**RESULTS**

This section provides indicative results on the merits of the CR concept. Merits are due to the complementary role of the wireless access networks of the CR infrastructure, and are obtained through the cooperation of the SNRM platforms, including the NES components that simulate the managed networks.

The service area is covered by three network technologies: UMTS, HIPERLAN2 and DVB-T. Figure 4 depicts the coverage. The UMTS and DVB networks cover 100 percent of the service area. HIPERLAN2 varies between service area regions. Each UMTS cell is equipped with two carriers. The target is to operate UMTS cells at no more than 75 percent loading on the downlink and 55 percent on the uplink, per carrier. Therefore, it can be considered that each cell can provide up to 110 voice connections, which means that up to 90 Erlangs of equivalent-to-voice traffic load can be supported with 1 percent blocking probability.

It can be considered that each cell can provide up to 110 voice connections, which means that up to 90 Erlangs of equivalent-to-voice traffic load can be supported with 1 percent blocking probability.

### Figure 4. Service area coverage.

<table>
<thead>
<tr>
<th>Service area</th>
<th>DVB coverage</th>
<th>Service area</th>
<th>DVB coverage</th>
<th>Service area</th>
<th>DVB coverage</th>
<th>Service area</th>
<th>DVB coverage</th>
<th>Service area</th>
<th>DVB coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100%</td>
<td>2</td>
<td>50%</td>
<td>3</td>
<td>10%</td>
<td>4</td>
<td>50%</td>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>90%</td>
<td>7</td>
<td>0%</td>
<td>8</td>
<td>100%</td>
<td>9</td>
<td>75%</td>
<td>10</td>
<td>0%</td>
</tr>
<tr>
<td>11</td>
<td>50%</td>
<td>12</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(e.g., loading factors at certain range and balance, segregation of services with different requirements from the same carrier). The second task is at a finer level of detail, since it configures the loading factors (interference patterns) per cell and carrier. It can rely on simple methods [2] or more elaborate schemes [11], depending on the level of detail of the input.

**HIPERLAN2 (Fig. 3c).** The first task is to partition the capacity between the uplink and downlink in each cell. The second task is the preservation of the appropriate C/I levels in each cell. In general, the C/I is an important indicator of the performance in each cell. The higher the C/I, the higher the physical layer mode that can be selected (by the link adaptation operation), and consequently the throughput that can be achieved in the cell. Therefore, the component manages the frequency selection and power control operations, which are means of achieving the target C/I levels. In principle, the maximum transmission power (of APs and terminals) that will be allowed by the power control operation in moderately loaded cells can be limited, since large C/I values (and the consequent high bit rate values) are not required. This enables more heavily loaded cells to more easily achieve the required high C/I values. Likewise, dense frequency reuse can be imposed to moderately loaded cells by prohibiting the use of certain frequencies in the respective APs. More heavily loaded cells can use the remaining, less utilized frequencies, and therefore achieve the required high C/I (bit rate) values.

DVB. In general, the DVB segment has an amount of bandwidth available and the ability to support a certain number of traffic classes (not related to TV programs) within given time zones. The objective of this component is to maximize the bandwidth assigned to these traffic classes, subject to capacity and profile constraints (e.g., provision of acceptable QoS levels and admission of appropriate services in the DVB domain).
In service area regions in which the alternatives are limited the CR infrastructure can improve the performance considerably. There can be roughly 10 percent improvement in the carried traffic in the areas in which there is only DVB and sparse HIPERLAN2 coverage. This improvement is adequate for maintaining all the classes at the highest possible QoS level. The improvement is more important in service area regions in which the CR infrastructure offers more capabilities. Specifically, the carried traffic can be increased by 30 percent in service area regions with more dense HIPERLAN2 coverage. This is achieved if the CR infrastructure can complement each UMTS cell by carrying an additional user bit rate of approximately 500 kb/s. This bit rate is achievable by HIPERLAN2 systems; however, it can require several HIPERLAN2 cells per UMTS cell, depending on the size of the latter and the demand distribution. Nevertheless, complementing a UMTS configuration, with small HIPERLAN2 (or wireless LAN in general) islands, is believed to be a cost-effective solution compared to that of deploying additional UMTS equipment.

At this stage, there can be sample explanations on the reconfigurations required from the managed networks. It is assumed that only area 6 faces the excessive load that derives from 9000 users. Figure 5b depicts the AP transmission powers in different areas of the HIPERLAN/2 network. Figures 5c and 5d depict the uplink and downlink interference levels (loading factors) in different areas of the UMTS network. As anticipated, areas that do not face excessive requirements can operate at lower levels. Therefore, the RMS-U will instruct the corresponding network segments to configure their control domains accordingly in order to assist area 6 to achieve the required capacity figures. Moreover, the CR concept enables the UMTS capacity to be used to maintain the high QoS levels offered to part of the demand. The rest of the demand will be maintained at high QoS levels by the other components of the CR infrastructure.

<table>
<thead>
<tr>
<th>Service</th>
<th>Voice</th>
<th>Videoconference</th>
<th>Interactive data</th>
<th>Streaming video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Networks</td>
<td>UMTS, HIPERLAN2</td>
<td>UMTS, HIPERLAN2</td>
<td>UMTS, HIPERLAN2, DVB</td>
<td>UMTS, HIPERLAN2, DVB</td>
</tr>
<tr>
<td>QoS levels description</td>
<td>Standard</td>
<td>Standard</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>QoS levels requirements</td>
<td>12.2 (60), 1%</td>
<td>64 (120), 1%</td>
<td>144 (240)</td>
<td>384 (480) 144 (240)</td>
</tr>
<tr>
<td>User classes</td>
<td>Gold</td>
<td>Economy</td>
<td>Gold</td>
<td>Gold</td>
</tr>
</tbody>
</table>

Table 1. Profile/policy information.

![Image](image_url)

**Figure 5.** a) Performance (carried traffic) of the CR infrastructure vs. different demand volumes; b) transmission power used in APs at different segments of the HIPERLAN/2 network; c) transmission power (dBm) in cells of different segments of the UMTS network; d) uplink interference in cells of different segments of the UMTS network.
Likewise, in the HIPERLAN2 network the cells serving the area with highest demand will obtain carriers with less interference and be allowed to operate at higher transmission powers.

In summary, the results from the SNRM application indicate the following. It is feasible to exploit the complementary role of the wireless systems in the CR infrastructure, and enable the achievement of higher capacity volumes and the provision of high QoS levels to user classes. Cost efficiency is boosted by the fact that it can be cheaper to rely on some service area segments and for some time zones, BRAN/WLAN technologies or DVB, and to gradually (at a slower pace) upgrade 2.5G or 3G infrastructures. Within each network the resources are managed (and the control domain is configured) to assist the areas that face excessive requirements. Due to the CR concept, the capacity gained can be utilized to offer high QoS levels to the largest possible part of the demand, while the rest is served in the same manner from other components of the infrastructure.

**CONCLUSIONS**

The convergence and interworking between different wireless access networks, in the context of a CR infrastructure, is a recent trend. The materialization of the concept of service delivery through a CR infrastructure requires new management concepts. In this article an SNRM platform for wireless systems operating in a CR context was introduced. Aspects addressed were platform requirements, high-level design, and the functionality of the main management entities. Moreover, results were presented indicating the merits of the CR concept in terms of increased capacity and QoS levels, and lower cost.

**REFERENCES**


**BIographies**

Panagiotis Demestichas received his diploma and Ph.D. degrees in electrical and computer engineering from National Technical University of Athens (NTUA). Since September 2002 he has been an assistant professor at the University of Piraeus, in the Department of Technology Education and Digital Systems. Since 1993 he has collaborated with the Telecommunications Laboratory of NTUA. His research interests include the design, management, and performance evaluation of mobile and broadband networks, software engineering, algorithms and complexity theory, and queuing theory.

Nikolaos Koutsourakis received his diploma in electrical and computer engineering from NTUA in 2001. Since September 2001 he has been a research associate at the Telecommunications Laboratory of NTUA. He is a post-graduate student in the M.B.A. course Technical-Economical Systems, co-organized by three major Greek universities (NTUA, University of Athens, and University of Piraeus). His research interests are in the areas of wireless communication networks and distributed middleware programming.

George Koukourakis received an electrical and computer engineering degree from NTUA, Greece, in July 2001. Since September 2001 he has been a research associate at the Telecommunications Laboratory at NTUA. He has worked in research projects in the context of the IST framework. He is the author of several scientific papers in the areas of mobile communications.

Konstantinos Tsigkaris received a diploma in electrical and computer engineering from NTUA in 2000. Currently he is a research associate at the Telecommunications Laboratory of the Department of Electrical and Computer Engineering of NTUA, working toward his Ph.D. degree in the area of wireless communications and broadband networks design.

Antonis Oikonomou received an electrical and computer engineering degree from NTUA, Greece, in 2000. Currently, he is a research engineer with the Telecommunications Laboratory of NTUA, also pursuing his Ph.D. in the area of 4G wireless communications. He has worked in national and international research projects in the context of the IST framework. He is the author of several scientific papers in the areas of mobile communications.

Vera Stavroulaki received a diploma in informatics from Athens University of Economics and Business in September 2000. Since September 2000 she has been a research associate at the Telecommunications Laboratory of the Department of Electrical and Computer Engineering of NTUA, working toward her Ph.D. degree in the area of wireless communications and broadband network design. She is currently involved in the European IST programs CREDO and ADAMANT.

Louisa Papadopoulou received her diploma in mining and metallurgical engineering and her M.Sc. in engineering-economic systems from NTUA in 1998 and 2002, respectively. Currently, she is a Ph.D. student and works as a research engineer in the Telecommunications Laboratory of NTUA. Her research interests are in the area of composite radio environments and related engineering-economic problems.

Michael E. Theologou received a degree in electrical engineering from Patras University and his Ph.D. degree from the School of Electrical Engineering and Computer Science of NTUA. Currently he is a professor in the School of Electrical Engineering and Computer Science of NTUA. His research interests are in the field of mobile and personal communications.

Guillaume Vivier graduated from the Ecole Nationale Supérieure des Communications (ENST), Paris, France in 1993. He first worked for Alcatel and joined Motorola Laboratories in 1998, where he is currently a senior staff engineer. His research interest includes digital communications, protocol design, and performance evaluation of wireless systems and reconfigurable systems.

Karm El-Khazen received his M.Sc. degree in electrical and computer engineering from the Georgia Institute of Technology and his engineering degree in networks and telecommunications from ENSEA, France in 1998. Currently, he is a senior research engineer at Motorola Laboratories, France. He is part of the European Communications Research Laboratory. His major research interests are in the areas of wireless and broadband communications, including management architectures in beyond 3G systems. He contributed to the European project MONASIDRE.