Mobility management in B3G networks: a middleware-based approach

Letian Rong, Manel Fredj, Valérie Issarny, Nikolaos Georgantas

Mobility Management in B3G Networks: a Middleware-based Approach

Letian Rong, Manel Fredj, Valérie Issarny and Nikolaos Georgantas
INRIA
Domaine de Voluceau, Rocquencourt, B.P. 105, 78153
Le Chesnay Cedex France
+33139635185
{Letian.Rong, Manel.Fredj, Valerie.Issany, Nikolaos.Georgantas}@inria.fr

[Position Paper]

ABSTRACT

The B3G (Beyond 3G) networking will enable mobile users to roam freely through heterogeneous networks on an all-IP platform. However, mobility handling in such an environment poses new challenges. Traditionally, mobility protocols such as Mobile IP, SIP and SCTP are used to manage mobility in B3G, but they require telecommunication companies to either modify existing network infrastructures or deploy central entities in the network core to handle mobility. This is not feasible in a fully distributed computing environment (e.g., P2P) and mobile ad hoc networks that are part of B3G networking. As an alternative, this paper introduces a middleware component with modularized functionalities to facilitate mobility management in a fully distributed B3G environment.

Keywords

B3G, mobility management, middleware, handoffs and service reconfiguration.

1. INTRODUCTION

Over the last decade, there has been a tremendous growth of wired and wireless networks (e.g., WLAN, Bluetooth and 3G), as well as proliferation of mobile devices (e.g., smart phones). The Beyond 3G (B3G) networking [1] aims to provide an all-IP platform for handling heterogeneous network access to cope with this phenomenon. This will allow users to roam freely through various networks using different access technologies and devices in a seamless manner. Also, as mobile devices are becoming more powerful, they can be utilized to host services.

In general, we view B3G networks as an aggregation of heterogeneous infrastructure-based and infrastructure-less (ad hoc) networks. In this context, we are designing a middleware-based approach to B3G networking so that access over B3G networks does not depend on the deployment of a central infrastructure in the core network and can effectively integrate ad hoc networks. This leads us to qualify the target network as infrastructure-less B3G network. In this paper, we more specifically concentrate on the issue of mobility in infrastructure-less B3G networks, which is challenged by the mobility of both service clients and providers. We first briefly survey the current trend in B3G networking (Section 2). We then outline the proposed middleware-based approach to B3G networking, as investigated in the PLASTIC project (Section 3). Mobility handling in that setting is further investigated, sketching requirements in the light of illustrative scenario (Section 4) and base middleware solution (Section 5). Finally, conclusions are drawn in Section 6.

ESSPE '07, September 4, 2007, Dubrovnik, Croatia
Copyright 2007 ACM ISBN 978-1-59593-798-8/07/09...$5.00

2. BACKGROUND

Mobility management in B3G networks is generally handled by means of network handoffs. Users are able to continuously access the same service through methods such as, access-point switching within the same network and data traffic redirection between two different networks. The former is called horizontal handoff and the latter vertical handoff.

Horizontal handoffs are often managed by the wireless technology at the data-link layer. On the other hand, vertical handoff is more complicated since it involves networks of different types. The most common approaches are based on the mobility protocols Mobile IPv4 (MIPv4) [3] and Mobile IPv6 (MIPv6) [4]. MIPv4 deploys central entities called the Home Agent (HA) and the Foreign Agent (FA) in the home network and the foreign network respectively to redirect packets from the provider to the mobile node when the node changes from the home network to the foreign network. MIPv6 then improves the performance of MIPv4 by removing the FA and allows the mobile node to directly receive packets from the provider once it has updated its new address with the HA. Numerous mobility management solutions have been proposed based on these two protocols [19]-[21]. However, the deployment of central entities such as HA and FA in networks is hardly feasible in an infrastructure-less B3G environment where mobile nodes communicate in an impromptu manner and networks are established on the fly. In addition, both MIPv4 and MIPv6 require modification to legacy networks as they are network layer based protocols.

In contrast, the Session Initiation Protocol (SIP) protocol [5] moves mobility management to the application layer and allows mobility solutions to be deployed in a network-agnostic manner [15]. However, it has similar feasibility problem as MIP since it uses a central SIP server in the network core to perform address update of mobile nodes [15].

Another popular protocol for vertical handoff is SCTP [6], which supports network switching through multi-homing. Multi-homing allows two mobile nodes to set up an association with multiple IP addresses for each node, and therefore allows applications to perform switchovers between different radio interfaces during vertical handoffs without interrupting data transfer. Unlike techniques based on MIP and SIP, SCTP-based solutions do not require the addition of central components to the existing network as handoffs are performed by mobile nodes themselves [22], [23]. However, SCTP is a transport layer solution. Thus, it is not practical to adopt in the real world since TCP and UDP are the most commonly used protocols for this layer.

Another class of mobility management solutions has emerged in recent years; these are based on service reconfiguration. Indeed, service reconfiguration enables service continuity by replacing the disconnected service with an equivalent service available in the client's environment [24]. For instance, both Gaia [9] and the SRR-WS (Suspend-Relocate-Resume for Web Services) framework [10], handle mobility by finding a service substitute, and putting it in a state allowing resumption of the interrupted client/service interaction. However, these approaches all take a centralized approach (i.e., use a single reconfiguration manager) to perform service substitution and reconfiguration. In addition, their usage in a B3G context has not been explored.

3. USE CASE SCENARIO AND REQUIREMENTS FOR DISTRIBUTED B3G

A use case scenario is discussed in this section to illustrate mobility handling in a B3G networking environment. We then present a set of requirements for mobility management.

3.1 Use Case Scenario

With significant advances in wireless technologies, as well as pervasive and wearable systems [11], [12], m-Health (i.e., mobile-Health) [13] is gradually becoming a reality. In the following scenario, doctors are able to diagnose, monitor and treat patients remotely through the use of medical sensors, which are integrated into the environment of patients [14]. This allows a doctor to attend a patient at the earliest time and continuously provide the patient with care remotely. However, in most cases, a patient is still required to be attended in person if his/her condition cannot be treated remotely. A doctor can either attend the patient by himself or direct an assistant or another doctor to treat the patient. They offer their services such as diagnoses as Web services through carried mobile devices.

In the scenario, an elderly person George suddenly becomes ill at home and his son Mike decides to contact their private General Practitioner (GP) Dr. Smith. At the moment, Dr. Smith is visiting a patient at his home and answers George's call using his PDA through the WLAN connection at the patient's home. Mike switches on all the available medical sensors and monitors in the environment so that Dr. Smith can diagnose George remotely with Mike's assistance. Mike uses his PDA to send data collected from the sensors and monitors to Dr. Smith through a local WLAN connection. In addition, videoconferencing is established between them to allow a more visual communication and assessment of the situation. After the initial survey of the patient, Dr. Smith determines that George needs further diagnosis and asks Mike to drive his father to his clinic where he will meet them.

Once Mike and George are outside, the WLAN connection becomes weak as the connected PDA moving out of its range. Therefore, the PDA switches the connection from WLAN to 3G to obtain a better outdoor coverage. Also, the videoconferencing between them is changed to teleconferencing to adjust to the lower bandwidth of the 3G connection.

On the way towards Dr. Smith's private clinic, however, George's condition starts to further deteriorate. Furthermore, the connection between Mike and Dr. Smith becomes inconsistent since both of them are now mobile at high speed due to driving. The inconsistence makes it difficult for them to teleconferencing and for Dr. Smith to receive medical updates about George's condition. At this time, Mike's PDA locates a similar diagnosis service offered by another doctor Dr. Franklin in a nearby hospital. Both Dr. Smith and he decide that it is better to take George to Dr. Franklin for further diagnosis. Mike uses his PDA to retrieve the diagnosis data of George that was collected by Dr. Smith and passes them to Dr. Franklin, who then uses the data in continuing evaluating the patient. For the next few days, while George is being treated in the hospital, Mike, Dr. Franklin and Dr. Smith continue to communicate with each other through exchanging messages and medical data (i.e., images and video clips).

3.2 Requirements for Mobility Management in a Distributed B3G Environment

The above use case scenario reveals several different mobility cases and suggests the corresponding handling mechanisms. First, when both Mike and Dr. Smith moved to the outdoor, their PDA switched from WLAN to 3G to provide them with better coverage. This should be handled through vertical handoff. Also, the videoconferencing between them was changed to teleconferencing to adjust to the resulting lower bandwidth. Next, after Mike's PDA found a hospital nearby, Mike decided that it was better to change doctors to deal with poor communication caused by mobility. Therefore, his PDA collected diagnosis data from Dr. Smith and passed them to Dr. Franklin of the hospital to continue the care of his father. Finally, messages and medical data were transferred among them during the time when George was hospitalised. However, the messages and data might not be delivered to the recipients immediately as their mobile devices could be out of reach due to unavailability of network coverage or switched off. Therefore, the information was temporary saved on the sender or in the network, before delivered asynchronously to the recipients. This can also be considered as a special mobility case since users become unreachable when they move out of range of network coverage.

It is evident that different mobility handling mechanisms are required to support comprehensive mobility management in B3G networks. We specifically consider four different types of mobility handling mechanisms:

1. It provides mobility through network handoffs: the handoff process may be triggered based on related context and user preferences. Since devices possibly communicate in an ad hoc manner, the handoff process must not rely on any central entities or network infrastructures to perform the handoff.

2. It provides mobility through service reconfiguration: since service providers can be as mobile as service users, it may be difficult to sustain connections between them regardless which network connection is used. Also, sometimes a service may be available in a single access network and switching to a different network will prevent users from accessing the service (i.e., a service running on a BT device). Service reconfiguration may then be used as an alternative approach to network handoff for handling mobility. It should find a service substitute, which is semantically or syntactically equivalent to the disconnecting service, and further transfer state to ensure consistency of the current service session.

3. It provides specific mobility management for streaming applications with multi-modality support (i.e., video, audio and text transcoding): multimedia streaming applications are extremely delay-sensitive and the mobility solution needs to ensure the smooth delivery of multimedia streams. Also, the mobility solution needs to adapt the stream according to the surrounding network conditions and possibly its modality (e.g., video to audio) when it is required. Furthermore, it should not rely on any central entities to perform streaming and related tasks (i.e., transcoding).

4. It provides specific mobility support by delivering messages and data to users, even when they (i.e., sender or receiver) are not reachable or offline for a period of time: This requires the mobility solution to monitor the presence of users and ensures the information is eventually delivered to the recipients in an asynchronous manner.

4. THE PLASTIC MIDDLEWARE FOR B3G NETWORKING

The PLASTIC middleware [2] aims to provide essential functions to enable B3G networking for service oriented applications. As shown in Fig. 1, the PLASTIC middleware is layered on top of a legacy networked software platform and below applications and services. Application and services invoke the middleware functionalities through a set of APIs. The middleware is divided into independent modules and therefore, applications and services can choose to use a subset of the middleware based on their specific requirements.

The middleware modules are divided into two layers, namely, the upper *PLASTIC Middleware Services* layer and the lower *PLASTIC Communication Middleware* layer. The *PLASTIC Communication Middleware* layer contains a set of three modules to provide the basic functionalities for communications and networking in B3G. The bottom *Multi-radio Device Management* module is in charge of sensing the available networks and retrieving their status, as well as controlling the network interfaces. The *Multi-radio Networking* module sits on top of the *Multi-radio Device Management* module and offers basic communication functionalities such as point-to-point and multicast communications. In addition, it implements a PLASTIC-specific addressing scheme (PLASTIC@) that uniquely identifies a PLASTIC-enabled device and consequently the services hosted by it. The PLASTIC@ is a unique ID which can be associated with one or more IP addresses that corresponds to the network interfaces/connections a mobile device currently has. Furthermore, *Multi-radio Networking* aims to select the best network connections for applications running on a device based on a global optimization algorithm [25]. *Web-service Oriented Communication* module enables SOAP-based interaction in a B3G network. Also, it supports multi-network routing, which allows access to services across networks, as long as there is a path among the heterogeneous networks between the client and target service, thanks to the development of network bridges.

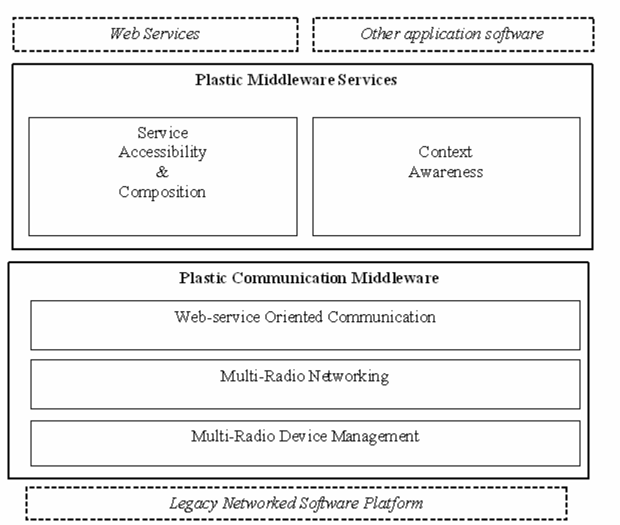

Figure 1. PLASTIC middleware

The upper *PLASTIC Middleware Services* layer contains advanced services related to distributed resource management. The related modules in this layer include *Service Accessibility & Composition* and *Context Awareness*. The *Service Accessibility & Composition* module enables Web services through service description, discovery and access in a context-aware manner. Also, it contains advanced functionalities such as service composition. Next, *Context Awareness* utilizes a context engine to allow efficient collection, storage and retrieval of context information in the B3G.

5. Mobility Management Module

Based on the scenarios and their analysis in Section 3, we propose to extend the PLASTIC middleware with a dedicated mobility management module to manage mobility. As shown in Fig. 2, the proposed module *Mobility Management* is inserted between the *PLASTIC Middleware Services* layer and the *PLASTIC Communication Middleware* layer. The *Mobility Management* module contains four sub-modules: *Network Handoff Management*, *Service Reconfiguration*, *Streaming Management* and *Message Delivery Management*. The functionalities of these sub-blocks correspond to the classified four types of mobility handling introduced in Section 3.3.

The *Network Handoff Management* module is responsible for managing mobility through the means of vertical and horizontal handoffs. It contains two sub-modules: *Multi-homing Mobility Management*, and *Horizontal Handoff Management*. The *Multi-homing Mobility Management* module is used to handle vertical handoffs through multi-homing (e.g., from 3G to BT). It utilizes the PLASTIC@ feature mentioned in Section 3.1 to perform IP address switching during multi-homing and aims to provide buffering to ensure smooth transitions. In contrast to SCTP-based approaches, multi-homing based on PLASTIC@ is a middleware layer solution and thus, is independent of underlying networks.

Horizontal Handoff Management allows horizontal handoffs to be initiated, thus decomposed into a number of sub-modules, each dedicated to a given wireless technology (e.g., WiFi and BT). Currently, WiFi does support horizontal handoff, however, the handoff is only performed in a ‘break-before-make’ fashion (i.e., hard handoff) and therefore, potentially causes packet losses in particular for delay-sensitive services. On the other hand, BT does not support automatic horizontal handoffs between BT devices. In general, these proposed modules aim to provide extra mobility methods for horizontal handoffs to complement standard approaches. These methods include support of smooth transition during horizontal handoffs and handoffs based on rules and contexts.

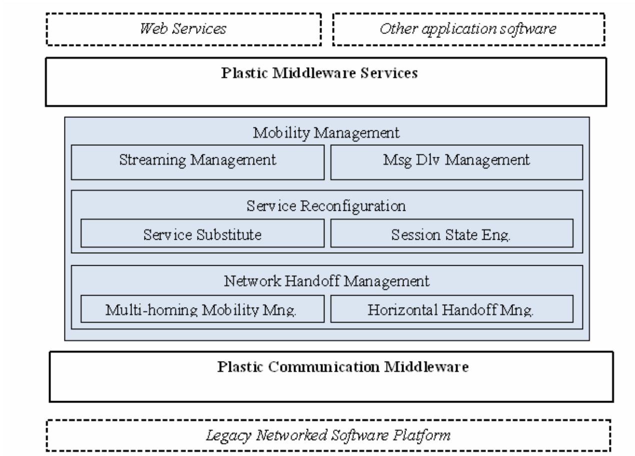

Figure 2. Extended PLASTIC middleware with the proposed mobility management module

All the sub-modules of *Network Handoff Management* use the underlying *Multi-Radio Device Management* module to retrieve network-related status and disconnect/connect network interfaces to perform handoffs. Furthermore, they utilize the *Context Awareness* module to obtain up to date context information (e.g., user location) in order to provide appropriate mobility management. In addition, *Network Handoff Management* should be extended to include other modules to support vertical and horizontal handoff solutions for other network types (e.g., WiMAX).

Service Reconfiguration intends to use service reconfiguration to facilitate mobility. Its functionalities are further categorized into two sub-blocks. *Service Substitution* stores a list of references to service substitutes, which can be used during the service reconfiguration process. The reference list can be either retrieved reactively or proactively. *Session State Engine* automatically saves the state of the current service in execution. It then adapts the saved state and transfers it to the service substitute during service reconfiguration to ensure service consistency.

In terms of interaction with other middleware modules, *Service Reconfiguration* utilizes the *Service Accessibility & Composition* module to perform service discovery during service reconfiguration. In addition, the *Context Awareness* module provides it with context information related to the management of mobility. It also uses the *Network Handoff Management* sub-module to switch between networks/APs (Access Points) when it is necessary during service reconfiguration (i.e., switching from WiFi to BT when trying to access a service substitute, which only exists on the nearby BT device).

Streaming Management and *Message Delivery Management* can be considered as the highest level sub-modules contained in the mobility management support and are tailored to specific application domains. *Streaming Management* is designed specifically to the needs of multimedia streaming applications. It aims to provide seamless streaming in a B3G environment. This requires the module to handle stream packet losses and high jitter caused by network congestions, as well as B3G-specific events such as horizontal and vertical handoffs. The former should consider existing solutions in distributed streaming [16], [17], while the latter can be handled using the *Network Handoff Management* sub-module to perform smooth handoffs. In addition, *Streaming Management* should support multi-modality based streaming to further improve the user experience. It can utilize the *Service Reconfiguration* sub-module to reconfigure to services representing different modalities, as well as using the upper layer *Service Accessibility & Composition* module to find transcoding services. Finally, it uses the *Context Awareness* module to provide adaptive multimedia streaming based on relevant context.

The last sub-module *Message Delivery Management* intends to support message and possibly data delivery in an asynchronous manner, so that the delivery can be made even if the sender and/or receiver are disconnected from the network for some time. It may in particular use the multi-network routing feature of the *Web-service Oriented Communication* module to determine possible network path to the intended destination. Furthermore, asynchronous data delivery proposals such as Haggle [18] are currently under investigation to determine their feasibility within the module.

The *Mobility Management* module aims to provide a set of simple and uniform APIs to ease mobility-aware service development. The APIs can be classified into two types based on their usage for mobility: transparent and explicit. The transparent type is used to hide lower level details of underlying B3G networks from application developers and provide automatic mobility management through the *Mobility Management* module. The behaviour of the automation can be either predefined within sub-modules (i.e., horizontal handoff API for BT based on signal strength of surrounding APs) or determined by a set of policies or cost-functions as inputs to the APIs. The feasibility of policy-based [7] or cost-function-based [8] approaches is currently under study. However, the application developers may want to explicitly perform a mobility-based function without passing application logic to the lower middleware. This is the case when the application logic is very specific and too complex to describe through rules or cost-functions. For example, a doctor who is travelling decides to pass the care responsibility of a patient to his mobile assistant because the 3G connection on his mobile device becomes too weak (i.e., service reconfiguration based mobility). Therefore, it is equally important to provide executable mobility functions explicitly through APIs.

6. Conclusion

In this paper, we have outlined the design of a middleware-layer mobility management module for infrastructure-less B3G networking environment. The module aims to handle mobility of both users and service providers through a set of functions. The two lower modules use multi-homing and service reconfiguration to handle mobility, while two higher modules provide mobility management for specific applications such as streaming and asynchronous message delivery. In addition, the mobility management modules provide a set of APIs, which allows application developers to choose to have mobility management handled either transparently or explicitly. To this end, we have designed the APIs and started implementing the two lower level sub-modules: *Network Handoff Management* and *Service Reconfiguration*.

7. ACKNOWLEDGMENTS

This work is part of the IST-6 project PLASTIC and has been funded by the European Commission, FP6 contract number 026955 (http://www.ist-plastic.org/).

8. REFERENCES

[1] Zahariadis T. and Doshi B. Applications and Services for the B3G/4G Era. *IEEE Wireless Communications.* Vol. 11, No. 5, October, 2004.

[2] Issarny V. et. al. PLASTIC Deliverable 3.1: Middleware Specification and Architecture. *PLASTIC Deliverable*. March, 2007. http://www.ist-plastic.org.

[3] Perkins C. IP Mobility Support for IPv4. *RFC 3344*. August, 2002.

[4] Johnson D., Perkins C. and Arkko J. Mobility Support in IPv6. *RFC 3775*. June, 2004.

[5] Rosenberg J. et al. SIP: Session Initiation Protocol. *RFC 3261*. June, 2002.

[6] Stewart R., Xie Q., and Morneault K. Stream Control Transmission Protocol. *RFC 2960*. October, 2000.

[7] Vidales P., Baliosian J. and Serrat J. Autonomic System for Mobility Support in 4G Networks. *IEEE Journal on Selected Areas in Communications*. Vol. 23, No. 12, 2005.

[8] Balasubramaniam S. and Indulska J. Vertical Handover Supporting Pervasive Computing in Future Wireless Networks. *Computer Communication Journal, Special Issue on 4G/Future Wireless Networks*. Vol. 27, No. 8, 2004.

[9] Roman M. et al. Gaia: A middleware platform for active spaces. *ACM SIGMOBILE Mobile Computing and Communications Rev.* Vol. 6, No. 4, 2002.

[10] Dorn C., and Dustdar. S, Achieving Web Service Continuity in Ubiquitous Mobile Networks the SRR-WS Framework. *UMICS 2006.* Luxemburg, June, 2006.

[11] Jovanov E., Lords A., Raskovic D., Cox P., Adhami R., and Andrasik F. Stress Monitoring using a Distributed Wireless Intelligent Sensor System. *IEEE Eng. Med. Biol. Mag.* Vol. 22, No. 3, May/June, 2003.

[12] Pattichis C. S., Kyriacou E., Voskarides S., Pattichis M. S., Istepanian R. S. H., and Schizas C. N. Wireless Telemedicine Systems : An overview. *IEEE Antennas Propagat. Mag*. Vol. 44, No. 2, April, 2002.

[13] Istepanian R. S. H., Jovanov E., Zhang Y. T. Guest Editorial Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity. *IEEE Transactions on Information Technology in Biomedicine*. Vol. 8, No. 2, December, 2004.

[14] Jovanov E., Price J., Raskovic D., Kavi K., Martin T., and Adhami R. Wireless Personal Area Networks in Telemedical Environment. *3rd International Conference on Information Technology in Biomedicine*. Virginia, USA, November, 2000.

[15] Banerjee N., Das S.K., Acharya A. Seamless SIP-Based Mobility for Multimedia Applications. *IEEE Network*. Vol. 20, No. 2, 2006.

[16] Drougas Y., Repantis T. and Kalogeraki V. Load Balancing Techniques for Distributed Stream Processing Applications in Overlay Environments. *ISORC'06.* Korea, April, 2006.

[17] Akyol E., Tekalp A. M., and Civanlar M. R. Adaptive Peer-To-Peer Video Streaming with Optimized Flexible Multiple Description Coding. *IEEE International Conference on Image Processing*. Atlanta, USA, October, 2006.

[18] Scott J., Hui P., Crowcroft J. and Diot C. Haggle: A Networking Architecture Designed Around Mobile Users. *IFIP WONS 2006*. January 18-20, Les Menuires, France.

[19] Valko A. Cellular IP: A New Approach to Internet Host Mobility. *ACM SIGCOMM Comp. Commun*. Rev. January, 1999.

[20] Bi Y., Song M. and Song J. Seamless Mobility Using Mobile IPv6. *IEE Mobility'05*. Guangzhou, China, November, 2005.

[21] Fan G. and Lu X. Novel Service Oriented Handoffs in Heterogeneous Wireless Networks. *QShine'05.* Orlando, USA, August, 2005.

[22] Ma L., Yu F., Leung V. and Randhawa T. A New Method to Support UMTS/WLAN Vertical Handover Using SCTP. *IEEE Wireless Communications.* Vol. 11, No. 4, August, 2004.

[23] Chen Y.-M., Lai M.-Y., Lin S.-C., Chang S.C. and Chung T.Y. SCTP-based Handoff Based on MIH Triggers Information in Campus Networks. *ICACT'06.* Phoenix Park, Korea, October, 2006.

[24] Fredj M., Zarras A., Georgantas N. and Issarny V. Connectivity Loss in Pervasive Computing Environments. In *Proceedings of the 4th MiNEMA Workshop*. Sintra, Portugal, July, 2006.

[25] Caporuscio M., Charlet D., Issarny V., and Navarra A. Energetic-Performance of Servive-oriented Multi-radio Networks : Issues & Perspectives. *WOSP*. Buenos Aires, Argentina, February, 2007.