

QoS and Security in 4G Networks

Xiaoming Fu¹, Dieter Hogrefe¹, Sathya Narayanan², Rene Soltwisch¹

¹Telematics Group,
University of Goettingen, Germany
{fu,hogrefe,soltwisch}@informatik.uni-goettingen.de

²Panasonic Information & Networking
Technologies Laboratory, USA
sathya@research.panasonic.com

Abstract

Future 4G mobile communication networks are expected to provide all IP-based services for heterogeneous wireless access technologies, assisted by mobile IP to provide seamless Internet access for mobile users. Two major challenges in developing such heterogeneous network infrastructure are QoS provisioning and security services for mobile users' communication flows. This paper proposes a new architectural view and methodologies for QoS and security support in 4G networks, which integrates QoS signaling with authentication, authorization and accounting (AAA) services to both guarantee the user applications' QoS requirements and achieve efficient authentication, authorization and key exchange.

1 Introduction

In the past decade, the telecommunications industry has witnessed an ever accelerated growth the usage of mobile communications. As a result, the mobile communications technology has evolved from the so-called second-generation (2G) technologies, GSM in Europe, IS-95 (CDMA) and IS-136 (TDMA) in USA, to the third generation (3G) technologies, UMTS/WCDMA in Europe and CDMA2000 in USA, being standardized by 3GPP and 3GPP2 (respectively), partnership projects between the governmental standards development organizations (SDO) of various countries.

Along with the standards development for providing voice service to mobile users, a group of standards to deliver data to the mobile users have evolved from both SDOs and industry. Systems and applications, such as iMode, the mobile Internet access system developed by NTT DoCoMo, and Short Message Service (SMS) for sending and receiving short text messages for mobile phone users, have been built and continue to be developed. The WAP (wireless application protocol) forum and more recently, the Open Mobile Alliance have also been developing applications for wireless networks.

In the wireless access field, Bluetooth was developed as a new cable replacement technology, which provides a short-range (~10m), low bit rate (1Mbps) access in the 2.4GHz spectrums. IEEE also developed a wireless LAN (WLAN) access family of protocol IEEE 802.11 including 802.11b (a 100m, 11Mbps access technology in the 2.4GHz spectrum), 802.11a, and 802.11g, as well as HiperLAN2 developed by ETSI. Nowadays, 802.11 has become one of the most popular and easy ways to provide wireless access for nomadic laptop users; first products of cellular phones that can access IEEE 802.11 base stations have recently been available in the market.

The focus of this paper is on fourth generation (4G) mobile networks. Even though a universal consensus on what is going to be 4G is not yet reached in the industry or the literature, there is a reasonable understanding of some characteristics of 4G mobile networks. Some of the accepted characteristics are:

- All-IP based network architecture;
- Higher bandwidth;
- Support for different access networks, including WLAN technologies like IEEE 802.11;
- Full integration of "hot spot" and "cellular";
- Support for multimedia applications.

In order to clarify our vision of 4G networks, we could imagine a staff A starts a voice over IP conversation with his boss B (who is at a remote site) on his way to the airport, through an access to the UMTS system. When he arrives at the airport, WLAN access becomes available and the conversation (connectivity) between A and B is expected to be seamlessly continued (upon necessary authentication of A's credential by the network) even with a different access technology and a different operator. Furthermore, data transmission over the wireless link may desire stronger protection and the conversation between A and B may desire certain QoS support from the network. This scenario indicates the integration of different characteristics of data transmission and data protection, and possible different approaches for quality assurance. As 4G is expected to be built on all-IP-based technologies, architectural considerations in IP layer become critical to enable

seamless interoperation among these technologies. While the IETF addresses the connectivity problems by its Mobile IP (MIP) protocols for both IPv4 and IPv6 networks [1,2], Quality of Service (QoS) and security insufficiencies are apparent: besides a simple IPsec support for MIP registration process [3], MIP mainly maintains the connectivity between a mobile node (MN) and its corresponding node (CN) while it is moving away from its home networks. It neither supports QoS nor stronger security between the MN and the network.

QoS mechanisms, including resource reservation (signaling), admission control and traffic control, allow multimedia applications to get certain quality guarantee e.g., on bandwidth and delay for its packets delivery. Providing QoS guarantees in 4G networks is a non-trivial issue where both QoS signaling across different networks and service differentiation between mobile flows will have to be addressed. On the other hand, before providing network access and allocating resources for an MN, the network needs to authenticate the MN's (or the mobile user's) credential. Furthermore, a security association needs to be established between the MN and the network to ensure data integrity and encryption. Thus, in order to achieve seamless handover, mobility, QoS and security technologies must be integrated.

The rest of this paper presents a new architecture, **Seamless Mobility with Security and QoS Support in 4G Networks (SeaSoS)**, to address these challenges, which integrates QoS signaling, AAA and key exchange into the 4G mobile networking infrastructure. We present our views on 4G networks design and analyze underlying fundamental problems in Section 2. Section 3 describes the SeaSoS architecture and how it addresses these problems. Section 4 compares SeaSoS with other approaches and outlines some future work.

2 A Basic Model for 4G Networks

QoS, security and mobility can be viewed as three different, indispensable aspects in 4G networks; however all are related to network nodes involving the controlling or the processing of IP packets for end-to-end flows between an MN and the CN. We show in this section how we view the 4G network infrastructure based on which we present the SeaSoS architecture in Section 3.

2.1 Two Planes: Functional Decomposition

Noting that an IP network element (such as a router) comprises of numerous functional components that cooperate to provide such desired service (such as, mobility, QoS and/or AAA – Authentication,

Authorization and Accounting), we identify these components in the SeaSoS architecture into two planes, namely the control plane and the data plane.

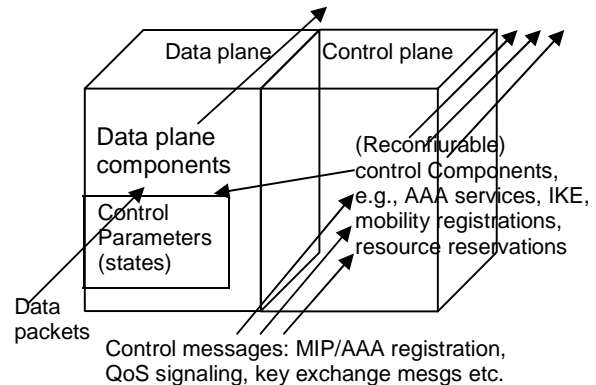


Fig. 1: The decomposition of control plane and data plane functionalities

Fig. 1 illustrates this method of flexible functional composition in 4G networks. As we are mainly concerned with network elements effectively at the network layer, we do not show a whole end-to-end communication picture through a whole OSI or TCP/IP stack. The control plane performs control related actions such as AAA, MIP registration, QoS signaling, installation/maintenance of traffic selectors and security associations, etc., while the data plane is responsible for data traffic behaviors (such as classification, scheduling and forwarding) for end-to-end traffic flows. Some components located in the control plane interact, through installing and maintaining certain control states for data plane, with data plane components in some network elements, such as access routers (ARs), IntServ [4] nodes or DiffServ [5] edge routers. However, not all control plane components need to exist in all network elements, and also not all network elements (e.g., AAA server) are involved with data plane functionalities. We refer these cases as path-decoupled control and other cases as path-coupled control.

We argue the separation and coordination of control plane and data plane is critical for seamless mobility with QoS and security support in 4G networks, with the reasons as follows. Per-flow or per-user level actions occur much less frequent than per-packet actions, while per-packet actions are part of critical forwarding behavior, which involves very few control actions (which are typically simply to read and enforce according the install state during forwarding data). Actually, this separation concept is not new – routing protocols have the similar abstraction together used with the traditional IP packet delivery, this abstraction is recently being investigated in

the IETF ForCES working group. However, we emphasize the three critical dimensions of future 4G networks: mobility, QoS and security, as well as other new emerging or replacement components might appear, integrated into a unified framework and allowing more extensibility for 4G networks design.

2.2 Two Modes of Operation

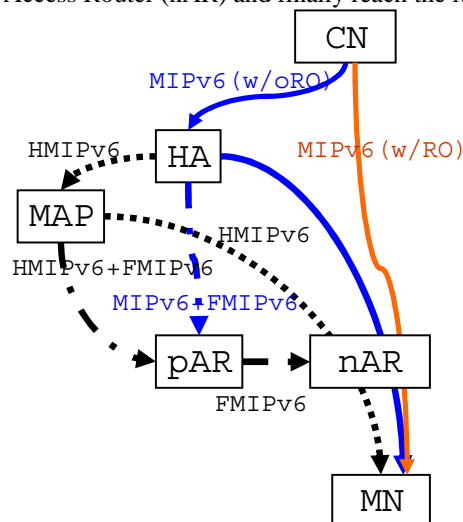
Besides the functional decomposition, we divide the operations that a 4G network infrastructure used in mobility scenarios into two categories: 1) end-to-end way control, which is related to authentication between the mobile device and the network, and enabling of forwarding end-to-end traffic. 1) the hop-by-hop way, mainly on hop-by-hop trust relationship and resource reservation setup.

2.3 Understanding Mobility, QoS and Security Problems in 4G Networks

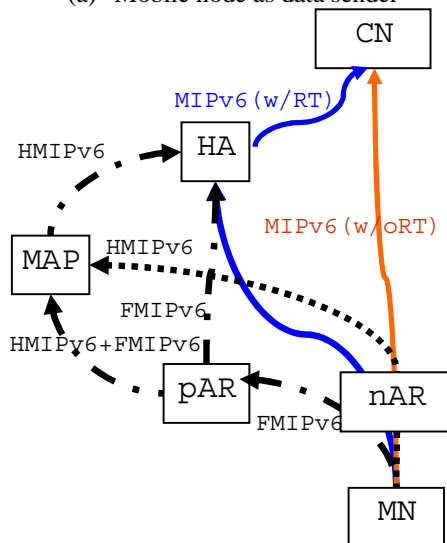
Mobility. Mobility involves both control plane and data plane. The control plane is mainly involved with path-decoupled, end-to-end way of mobility registrations, while data plane concerns mobility-enabled routing for data flows into and from an MN while it moves between different locations. The data plane behavior is achieved by installing/changing certain binding caches upon certain control plane information exchange (e.g., the binding update/acknowledge procedure in MIP). In fact, although MIP does not change the traditional IP routing table, when the MN is away from home and changes its location, associated with its fixed home address information, routing information is added in certain data processing and/or forwarding entities such as mobility agents (e.g., home agent and foreign agent) and systems themselves upon successful MIP registrations. Localized mobility solutions such as fast handover for Mobile IPv6 (FMIPv6) [13] and Hierarchical Mobile IPv6 (HMIPv6) [14] make this a little bit more complicated. Fig. 2 illustrates this issue in various MIPv6 cases.

As shown in Fig. 2(a), after different combinations of MIP registrations an MN can receive data flows along different paths sending from the CN. For example, after a MIPv6 with route optimization (MIPv6w/RO) registration, data flows traverse along normal IP routing path within bothering mobility agents. However, if a MIPv6 without route optimization (MIPv6w/oRO) registration is enforced, data flows can either traverse through the home agent (HA) towards MN directly (the normal case), or traverse through the HA and the mobility anchor point (MAP) introduced in HMIPv6.

Furthermore, when FMIPv6 is applied, the path can be more complicated by the way of further tunneling data packets from the Previous Access Router (pAR) through the New Access Router (nAR) and finally reach the MN.



(a) Mobile node as data sender



(b) Mobile node as data receiver

Fig. 2: A data plane view of an MN's flow

Similarly, Fig. 2(b) demonstrates the various potential data paths along which flows sent by the MN traverse, including the case after a MIPv6 registration with or without reverse tunneling (MIPv6w/RT or MIPv6w/oRT), or combined with FMIPv6 and/or HMIPv6. A more detailed description of these scenarios is provided in our prior work [8].

QoS. QoS provisioning also comprises data plane (mainly traffic control e.g., classification and scheduling)

and control plane (mainly admission control and QoS signaling) functions. Follow the above exploration of mobility problems, we can identify the fundamental difference of QoS provisioning in all-IP 4G mobile networks from a traditional, wired or wireless IP networks: whereas its resource control mechanisms can be similar to that of traditional networks, changing a location during the lifetime of a data flow introduces changed data path, thus requiring identifying the new path and installing new resource control parameters via path-coupled QoS signaling. Hence, a problem is how to apply any QoS signaling mechanism to achieve end-to-end resource setup in mobility scenarios. The current QoS signaling protocol, RSVP [7], exhibits lack of intrinsic architectural flexibility in adapting to mobility requirements. Difficulties arise, for example, because of its inability to adapt to the introduction of mobility routing in the data plane encountered in 4G networks, which results in either too complicated solutions or simply being unable to satisfy the needs. Over the years, research efforts have been made to address this (e.g., [8,19,20,21]) however it remains still an open issue.

Security. Security in 4G networks mainly involves authentication, confidentiality, integrity, and authorization for the access of network connectivity and QoS resources for the MN's flows. Firstly, the MN needs to prove authorization and authenticate itself while roaming to a new provider's network. AAA protocols (such as Radius, COPS or Diameter [10]) provide a framework for such support especially for control plane functions (including key establishment between the MN and AR, authenticating the MN with AAA server(s), and installing security policies in the MN or ARs' data plane such as encryption, decryption, and filtering), but they are not well suited for mobility scenarios. There needs to an efficient, scalable approach to address this. The Extensible Authentication Protocol (EAP) [6], a recently developed IETF protocol, provides a flexible framework for extensible network access authentication and potentially could be useful. Secondly, when QoS is concerned, QoS requests needs to be integrity-protected, and moreover, before allocating QoS resources for an MN's flow, authorization needs to be performed to avoid denial of service attacks. This requires a hop-by-hop way of dynamic key establishment between QoS-aware entities to be signaled on. Finally, most security concerns in this paper lie in network layer functions: although security can also be provided by higher layers above the network layer (for example TLS [15] provides privacy and data integrity between two communicating applications), our study mostly lies on mobility in the sense of network layer information exchange for mobile devices.

3 The SeaSoS Architecture

Reification of network architectures for support of QoS provisioning and security in 4G networks calls for new sights of dealing with the complexity of visualizing and architecting networks. Based on the network model described in Section 2, we present an architecture for **Seamless** Mobility with **Security** and **QoS** Support (SeaSoS), that integrates mobility schemes with QoS and security measures, and discuss the main issues toward realizing SeaSoS architecture.

SeaSoS differs from priori work in two main aspects: 1) it provides a distinct abstraction on functional separation and coordination of various involved network elements, which facilitates the network architects with a systematic exploration of the network design space. 2) SeaSoS allows network operators and end users to modify network attributes using dynamic plug-ins (e.g., replacing a mobility management protocol) or by re-configuring existing network services (e.g., adjusting the configuration parameters of traffic selector in an IPsec architecture). For example, SeaSoS allows MNs to re-configure their protocol stacks (e.g., from HMIPv6 to standard MIPv6 for mobility support, from Radius to Diameter or COPS for the AAA procedure) in order to dynamically interact with heterogeneous wireless access networks, and choose a certain QoS signaling protocol (such as RSVP or NSIS-QoS [9]) for their end-to-end applications. In one word, SeaSoS identifies the critical infrastructure of future 4G networks, as well as other new emerging or replacement components might appear, integrated into a unified framework and allowing an efficient, scalable and extensible network design for 4G networks. Note a network element may contain zero¹, one or many of data plane and control plane components in it.

As an example of basic SeaSoS operation, we use MIPv6, RSVP, AAA and EAP together in achieving a seamless handover in 4G networks. As shown in the Message Sequence Chart (MSC) in Fig. 3, we extend the method proposed in [22], namely apply EAP to perform the mutual authentication between the MN and access network, combined with AAA registration and extended with QoS and mobility support. Note the security association between the MN and the network is not directly transferred over the wireless interface, to avoid malicious nodes to obtain or modify it. As the MIP registration is also an end-to-end way operation, we

¹ Here we do not regard normal IP routing as part of control plane or data plane, therefore if a node only forwards normal IP packets it is transparent to other components related to mobility, QoS and security.

extend this approach to support efficient MIP registration. These transactions are shown in step 1-17. In addition, once a mobility registration takes place in the HA, a QoS signaling process can start for the flow destined to the MN. Here we use RSVP Path-Resv two-way signaling (step 18-23) but different from traditional RSVP, we use the combination of MN's permanent address (i.e., the home address) and the flow label as the unique identifier, which avoids a double reservation problem as identified in [8]. In order to prevent denial of service of QoS resources, we could apply an RSVP Integrity object [12] to the Path/Resv messages. Before applying this authenticated RSVP signaling procedure one may create a chain of trust relationship (security associations) along the RSVP nodes through the use of ISAKMP [17] with RSVP DOI [18] in the key exchange protocol IKE [16].

MSC - Basic SeaSoS operation (MIPw/oRO, CN->MN flows)

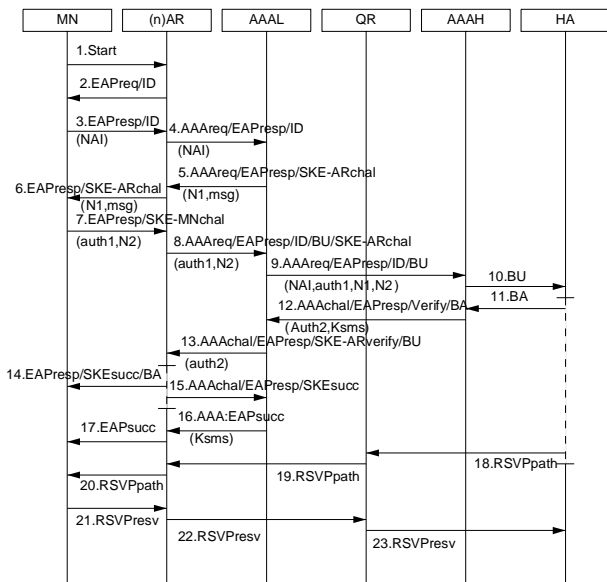
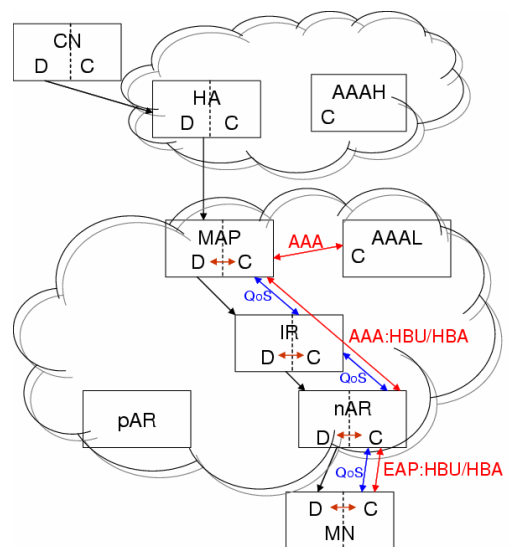


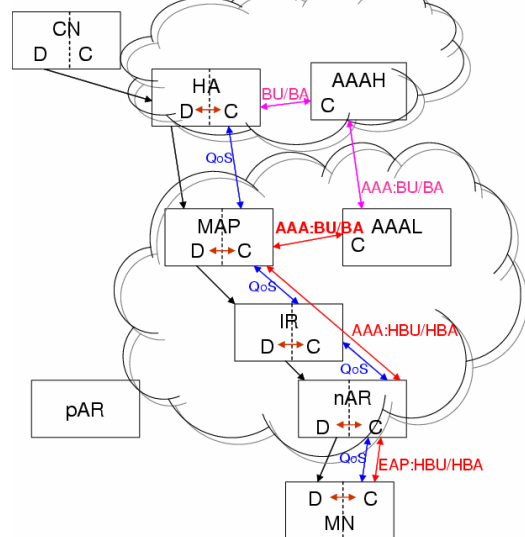
Fig. 3: Example of SeaSoS basic operation (control plane)

To further elaborate SeaSoS concepts, let's assume HMIPv6+MIPv6 is now the replacing mobility scheme. The simplest scenario can be that the MN just moves inside a MAP domain. As shown in Fig. 4(a), one can integrate the AAA procedure together with the HMIPv6 registration procedure. For example, if we use Diameter as the AAA protocol, we could encapsulate the HMIPv6 Binding Update (HBU) and Acknowledgement (HBA) messages inside the AAA request the MN control plane, similar to the Diameter Mobile IPv4 application [11]. This is rather simple, and after the MAP control plane receives this message, it forwards only the AAA request part to the local AAA server (AAAL); the latter authenticates the MN and if succeeds, returns a positive

AAA response to the MAP. Afterwards, MAP control plane changes its binding cache (i.e., mobility routing information for CN→MN traffic) in the data plane. Then the MAP can start two procedures without distinction of subsequence: 1) forward the AAA request with HBA through the nAR towards the MN (while traversing nAR, the control plane in nAR installs a traffic selector in the data plane for CN→MN traffic), 2) initialize a QoS signaling process towards the MN. Note 1) and 2) theoretically can be further merged but this increases the complexity of implementation.



(a) Mobile node moves inside a MAP domain



(b) Mobile node moves to a new MAP domain

Fig. 4: SeaSoS in HMIPv6 (flows destined to MN)

An inter-domain handover is similar (shown in Fig. 4(b)).

The difference lies in when the MAP determines it is a request to this domain, it initializes a AAA process combined with a global MIP registration (AAA:BU/BA). When the home AAA server (AAA:HA) accepts the request, a global binding cache is changed in HA's data plane; the HA can further initialize a QoS signaling process towards the MN. In both cases, we can see the handover process incorporates support of QoS, authentication and authorization for MN's flow in the HMIPv6 registration.

4 Summary and Future Work

There have been a few investigations on different aspects on QoS and security in 4G networks, notably MobyDick [19], SeQoMo [20], FCAR [21], and W-SKE [22]. We compare SeaSoS with them in Table 1.

Table 1: Comparison of SeaSoS with other approaches

Approach	Mobility support	QoS signaling	Security	Key exchange
MobyDick	MIPv6+ HMIPv6	Implicit session signaling	COPS/Diameter	No
SeQoMo	HMIPv6	QoS Cond.-Handoff	Diameter	No
FCAR	MIPv4	Changed RSVP	No	No
W-SKE	No	No	EAP+ Radius	Yes
SeaSoS	Any combination of IP mobility	Changed RSVP or NSIS-QoS	EAP+ any AAA	Yes

From this table we can see, different from these approaches, which mostly focus on functional aspects and optimization for certain circumstances, SeaSoS introduces the concept of supporting seamless mobility with and QoS mechanisms and security architectural components, which allows dynamic replacing/switching mobility management protocols and re-configuring existing network services in a secure way, and examines how they can be integrated universally to build an environment supporting 4G service requirements.

Due to the space limitation we can only sketch the key SeaSoS concepts in this paper. Towards the realization of the SeaSoS architecture, there are a number of issues to be resolved. A key issue is how to tradeoff between efficiency and security, especially when coordinating different control plane components. We are currently developing details of the proposed concepts in the context of seamless inter-domain mobility, and will validate the design through simulations and performance evaluations in a mobile IPv6 environment.

References

- [1] C. Perkins, IP Mobility Support for IPv4, RFC 3344, Aug. 2002.
- [2] D. Johnson, C. Perkins, and J. Arkko, Mobility Support in IPv6, RFC 3775, June 2004.
- [3] J. Arkko, V. Devarapalli, and F. Dupont, Using IPsec to Protect Mobile IPv6 Signaling between Mobile Nodes and Home Agents, RFC 3776, June 2004.
- [4] B. Braden, D. Clark, and S. Shenker, Integrated Services in the Internet Architecture: an Overview, RFC 1633, June 1994.
- [5] S. Blake, D. Black, and *et al.*, An Architecture for Differentiated Services, RFC 2475, Dec. 1998.
- [6] B. Aboba, L. Blunk, and *et al.*, Extensible Authentication Protocol (EAP), RFC3748, June 2004.
- [7] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification, RFC 2205, Sept. 1997.
- [8] X. Fu, H. Schulzrinne, and H. Tschofenig, Mobility Support in NSIS, Internet draft, June 2003.
- [9] S. Van den Bosch, G. Karagiannis, and A. McDonald, NSLP for Quality-of-Service Signaling, Internet draft, May 2004.
- [10] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, and J. Arkko, Diameter Base Protocol, RFC 3588, Sept. 2003.
- [11] P. Calhoun, T. Johansson, and *et al.*, Diameter Mobile IPv4 Application, Internet draft, July 2004.
- [12] F. Baker, B. Lindell, and M. Talwar, RSVP Cryptographic Authentication, RFC 2747, Jan. 2000.
- [13] H. Soliman, C. Castelluccia, K. El-Malki, and L. Bellier, Hierarchical Mobile IPv6 Mobility Management (HMIPv6), Internet draft, June 2004.
- [14] R. Koodli, Fast Handovers for Mobile IPv6, Internet draft, July 2004.
- [15] T. Dierks and C. Allen, The TLS Protocol Version 1, RFC 2246, Jan. 1999.
- [16] D. Harkins and D. Carrel, The Internet Key Exchange (IKE), RFC 2409, Nov. 1998.
- [17] D. Maughan, M. Schertler, M. Schneider, and J. Turner, Internet Security Association and Key Management Protocol (ISAKMP), RFC 2408, Nov. 1998.
- [18] H. Tschofenig and H. Schulzrinne, RSVP Domain of Interpretation for ISAKMP, Internet draft, Oct. 2003.
- [19] V. Marques, R.L. Aguiar, and *et al.*, An IP-based QoS Architecture for 4G Operator Scenarios, IEEE Wireless Communications, 10(3): 54-62, June 2003.
- [20] X. Fu, T. Chen, A. Festag, H. Karl, G. Schaefer, and C. Fan, Secure, QoS-Enabled Mobility Support for IP-based Networks, IPCN'2003, Dec. 2003.
- [21] S.-C. Lo, G. Lee, W.-T. Chen, and J.-C. Liu, Architecture for Mobility and QoS Support in All-IP Wireless Networks, IEEE JSAC, 22(4): 691-705, May 2004.
- [22] L. Salgarelli, M. Buddhikot, J. Garay, S. Patel, and S. Miller, Emerging Authentication and Key Distribution in Wireless IP Networks, IEEE Wireless Communications, 10(6): 52-61, Dec 2003.