A Network Mobility Management Scheme for Fast QoS Handover

Cheng-Wei Lee
Dept. of Information Management,
National Taiwan University
No.1, Sec. 4, Roosevelt Rd.,
Taipei City 106, Taiwan
+886-2-33661200
jameslee@net.edu.tw

Meng Chang Chen
Institute of Information Science,
Academia Sinica
128 Academia Road, Section 2,
Nankang, Taipei 115, Taiwan
+886-2-27883799#1802
mcc@iis.sinica.edu.tw

Yeali S. Sun
Dept. of Information Management,
National Taiwan University
No.1, Sec. 4, Roosevelt Rd.,
Taipei City 106, Taiwan
+886-2-33661200
sunny@im.ntu.edu.tw

ABSTRACT
The evolution of wireless access technologies has led to a new era of Mobile Internet. Network mobility, which considers the mobility of an entire network, is particularly suitable for mobile environments. Efficient network mobility QoS-handover design is essential to meet QoS requirements for real-time communication. In this paper, we propose a cross-layer hierarchical network mobility architecture and protocol, called HiMIP-NEMO, which is designed for all-IP networks. HiMIP-NEMO optimizes the routing between a mobile network node and the correspondent node, and supports fast QoS provisioning in the network mobility service domain. The simulation results demonstrate that Hi-MIP-NEMO reduces handover latency and packet loss, and supports high velocity vehicles.

Categories and Subject Descriptors
C.2.1 [**Network Architecture and Design**]: Wireless communication; C.2.3 [**Network Operations**]: Network management; D.4.8 [**Performance**]: Simulation.

General Terms
Management, Performance, Design.

Keywords
Mobile IPv6, Network Mobility, QoS Handover

1. INTRODUCTION
A concept called network mobility has become popular for supporting such high-mobility mobile networks. Network mobility refers to the mobility of an entire network that changes its point of attachment to the Internet as one unit, and all data packets sent to and from the mobile network are transmitted via one or more designated mobile routers (MR). Figure 1 shows an example that the mobile hosts in the mobile network (called mobile network node, or MNN) are immobile inside a moving vehicle which is equipped with an MR and provides network access service to the MNN. The vehicle usually moves within the wide area wireless network coverage of a single operator, e.g., Operator 1, and rarely crosses the boundary. The wide area wireless network can be an IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX) network or an IP-based cellular system. The major advantage of network mobility is that it reduces the overhead of handovers as well as the power consumption of the MNN.

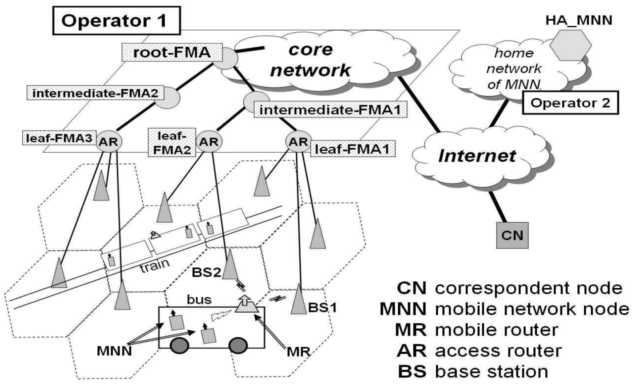

Figure 1. Network Mobility architecture

One basic requirement is Global Internet Reachability, which means that no matter where a mobile host roams, it is always reachable via its unique identifier, e.g., the home address (HoA) in Mobile IPv6 [1]. Mobile IPv6 is a popular method for supporting terminal mobility, but it does not explicitly support network mobility. Therefore, the Internet Engineering Task Force (IETF) extends Mobile IPv6 and publishes the NEMO basic support protocol [2] (NEMO-bs). By establishing a bi-directional tunnel between the home agent of the MR (HA_{MR}) and the MR, which traffic to/from the MNN must go through, NEMO-bs achieves transparency for the MNNs. However, the tunnel introduces packet header overhead and a suboptimal routing path, both of which cause unnecessary end-to-end transmission delay and jitter. In addition, NEMO-bs has inherent weaknesses that impact on handover.

Handover is the process of redirecting the traffic of on-going session(s) from an access point (i.e., BS, in this paper) to another. Since supporting real-time multimedia applications, for example, VoIP, is expected to be a promising service in mobile network, QoS-handover becomes a more important and complex process. It is essential that the resources needed for the existing services must be reserved in time when the mobile network is moving across the coverage area of BSs to ensure a successful QoS-handover, not just network connectivity re-establishment.

Traditional layered protocol reference model can no longer be efficient to support fast QoS-handover for network mobility service, where individual protocol layers are independently designed and functionally they do not cooperate. For example, during handover, the radio link attachment procedure to a BS in a new subnet may have been completed while the network-layer handover procedure is still waiting for the reconfiguration of mobile nodes' IP addresses. This inefficiency dictates a combined cross-layer mobility management and resource allocation design to reduce latency and packet loss during handovers.

In this paper, we propose a QoS-integrated network mobility architecture and protocol, called HiMIP-NEMO, which is designed for all-IP networks. When a vehicle travels in the network mobility service domain, the hierarchical architecture and the cross-layer designed protocols provide fast QoS provisioning for the mobile network in the vehicle. The function of the MR is simplified, and an MNN can use its applications and even security mechanism such as IPsec without modification in the mobile network.

2. RELATED WORK

To achieve route optimization, the MR in Prefix Scope Binding Update (PSBU) [3] send aggregate binding updates (BUs) for the mobile network to every correspondent nodes (CN), while in MIRON [4] and ROB [6], the MR helps the MNN to configure a geographically meaningful care-of address (CoA).

The handover performance of NEMO-bs, PSBU, ROB, MIRON, SIP-NEMO [7] and SIP-NMG [5] depends on how fast the new care-of address/prefix configuration and re-registration processes of the MR and the MNNs can be completed. According to [8], duplicate address detection must be run before the newly configured CoA can be accepted, even when the address is provided by the DHCPv6 server. In addition, subsequent re-registrations are time-consuming processes that degrade the handover performance. On the network level, fast handover for Mobile IPv6 [9] is designed to operate with the assistance of the link layer for reducing handoff latency and packet loss, and can be applied on NEMO-bs (FNEMO-bs). However, FNEMO-bs suffers from handover trigger timing, multiple tunnels and packet transmission delay fluctuation problems.

Current resource reservation protocols, such as RSVP [10], MRSVP [11] and HMRSVP [12] can only be applied to end-to-end reservation. NEMOR [13] and the BSR scheme [14] are proposed to support QoS for NEMO-bs. NEMOR uses a generic signaling protocol called Next Step In Signaling (NSIS) [15] to exploit DiffServ on the bi-directional tunnel between the MR and the HA_{MR} and IntServ between the HA_{MR} and the CNs. On the other hand, BSR adopts the concept of HMRSVP to manage network mobility with QoS guarantee. However, the MRs in NEMOR and BSR can only execute re-reservation process after regaining network connectivity under the BSs in new subnet, which degrade the handover performance. In addition, they cannot cooperate with FNEMO-bs that is designed for fast regaining network connectivity.

3. HiMIP-NEMO ARCHITECTURE AND PROTOCOL

Assume that an MR gets a new network prefix from its egress interface at a new location using prefix delegation protocol, and sends router advertisements (RA_{MR}) to the mobile network. Consider an MNN adopting Mobile IPv6 and having an HoA. When attached to the MR, it listens for RA_{MR}, and configures a geographically meaningful CoA which will be used to optimize the routing of packets. Once connected, the MNN may for example place a VoIP call.

3.1 System Overview

To achieve fast and QoS-guaranteed handover, we propose the use of Foreign Mobility Agent (FMA) in a hierarchical routing architecture which is typical in access networks. Using Figure 1 as an example, FMAs are functional modules residing in the routers of the network mobility service domain and are connected by a high speed wired network to facilitate their fast communications to support QoS-handover. FMA is responsible of reserving resources and allocate new resources for mobile networks once handover happened. A root-FMA is the edge router of the service domain. All other FMAs between the root-FMA and the leaf-FMA are called intermediate-FMAs. A leaf-FMA resides in access router (AR) having wired connections to several BSs, which can form one or many subnets. It is also responsible for managing radio channel resources of its child BSs.

The switching-FMA (SWF) refers to the cross-point of the old and new transmission paths of a handover instance. For example, if the MR in Figure 1 is in the subnet of BS1 and plans to move to the neighboring subnet of BS2, a handover occurs and the intermediate-FMA1 becomes the SWF. A main function of FMA is to maintain routing table for all the MRs attached to their child BSs.

In this paper, we focus on the handover procedure between BSs in different subnets. In the proposed HiMIP-NEMO architecture, a QoS-incorporated registration process is performed whenever an MR enters the network mobility service domain. Furthermore, we distinguish two types of handover in the HiMIP-NEMO operations: proactive and reactive to speed up the handover process. Two layer 2 messages from the wireless access network are used to trigger HiMIP-NEMO operations. When attaching to a BS, an MR will send a registration message, MRreg, containing its information to the BS. When the MR boots up in or enter the network mobility service domain, the MRreg will lead to a HiMIP-NEMO registration process.

The other layer 2 message is the handover indication (HOind) message which is sent by an MR when it is about to disconnect from the current serving BS and hand over to a target BS. An HOind will be sent only when the layer 2 pre-handover negotiations is successful. When receives the HOind message, a HiMIP-NEMO defined layer 3 message, **MRHOnotify**, will be sent by the serving BS to the target BS. During this message forwarding, based on the prefixes of the serving and target BSs, a route optimization is performed and the two BSs' nearest common ancestor (i.e. the SWF) is found in the hierarchical Mobile IP backhaul network. This refers to HiMIPv6-NEMO proactive handover. In addition, when abnormal condition happened, which causes failed proactive handover, a HiMIPv6-NEMO reactive handover will take over to recover the on-going sessions.

The REG-REQ and MOB_HO-IND messages defined in the IEEE 802.16e standard [16] have similar functionalities as MRreg and HOind used here, respectively. But they do not provide similar functions here whereby the BSs and FMAs in the hierarchical Mobile IP backhaul network cooperate to speed up optimal routing and resource allocation in efficient QoS-handover.

3.2 The QoS-incorporated Registration Process

The messages exchanged during the registration process are shown in Figure 2. In HiMIP-NEMO, during MR's registration, an extra field is defined in the MRreg to carry the prefix of the mobile network ($prefix_{MR}$). The MR then aggregates QoS requirements of the MNNs and sends the information to the BS. If the serving BS does not have QoS and handover related information of the MR, it send a HiMIP-NEMO defined layer 3 message, **MRinfo**, to its leaf-FMA containing the MAC address (MAC_MR), the $prefix_{MR}$, and the QoS parameters of the MR (QoS_{MR}).

The leaf-FMA first checks if there is enough radio channel resource on the BS. If not, a HiMIP-NEMO defined message, **MRinfo_reply**, is sent to the BS, rejecting the entrance of the MR. Otherwise, the leaf-FMA searches its MR list. If not found, the leaf-FMA creates a new record for the MR and records all the information into it including the QoS requirements. It then sends a HiMIP-NEMO defined message, **newMRquery**, to the root-FMA (step 2.) with the MAC_MR, $prefix_{MR}$, and QoS_{MR}. In the meantime, each intermediate-FMA on the path will also search its MR list for a record of the MR. If not found, a new one is created and the information in the newMRquery is copied into the record. When the root-FMA receives the message and does not find a record on its MR list, it replies a newMRreply with the MAC_MR, and the newMR and the QoS_reservation_confirm (QoSrc) fields set to "true" to the leaf-FMA (step 3).

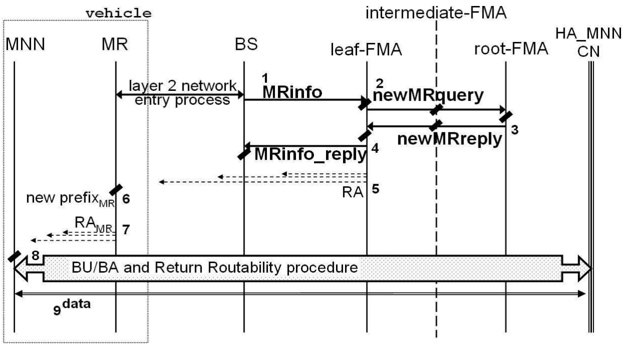

Figure 2. The QoS-incorporated registration process

On the way of passing the newMRreply from the root-FMA back to the leaf-FMA, if an intermediate-FMA reads the true values of the newMR and the QoSrc field, it will retrieve the QoS_{MR} information from its MR list and make corresponding QoS reservation. If the reservation succeeds, it simply forwards the newMRreply to the next FMA; otherwise, it sets the QoSrc field to false in the newMRreply to terminate further unnecessary reservations. If the leaf-FMA receives the newMRreply with the newMR and the QoSrc field set to true, it means that the MR is new in this domain and that all the intermediate nodes in the HiMIP-NEMO backhaul network from the leaf-FMA to the root-FMA are now ready to serve the MR and its following handovers. The leaf-FMA then sends a MRinfo_reply to the BS (step 4), accepting the entrance of the MR. For end-to-end QoS guarantees, the root-FMA may need to follow resource reservation protocols on behalf of the mobile network to reserve resources outside its domain to the CN; however, this procedure is beyond the scope of this paper.

After the registration, the MR receives router advertisement (RA) from the leaf-FMA (step 5), gets a new $prefix_{MR}$ and broadcasts RA_{MR} (step 6). The MNN then performs the normal Mobile IPv6 operations (steps 8), and starts to communicating with CN (step 9).

3.3 QoS-handover protocols

Figure 3 shows the procedure and the messages exchanged during a proactive handover. A layer 2 message, HOind, is sent by an MR when it is about to disconnect from the serving BS and hand over to a target BS (step 1). An extra field is defined in the HOind message to carry the $prefix_{MR}$. When receives the HOind, the serving BS sends an **MRHOnotify** to the target BS (step 2), carrying the MAC_MR, the $prefix_{MR}$ and the QoS_{MR}. Note that on the way of forwarding the MRHOnotify, the first FMA that finds both the serving and target BSs are its children, becomes the SWF (step 3). An FMA under the SWF on the old path, i.e., the ancestor of the serving BS but not the target BS, will execute an expiration function to release prior QoS reservations. The expiration function should be carefully calibrated to alleviate the ping-pong effect.

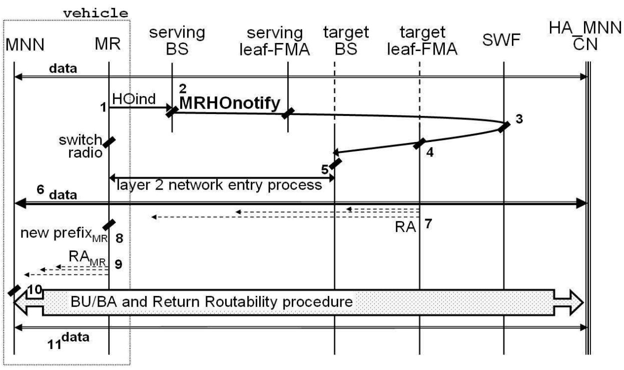

Figure 3. The QoS-incorporated proactive handover

The SWF creates a routing rule for this handover instance in its routing table to redirect packets whose prefix is the (old) $prefix_{MR}$ to the target BS. In step 4, each FMA on the path from the SWF towards the target BS will create a new record, and a routing rule for the mobile network as the SWF does. In addition, the leaf-FMA checks whether there is enough radio channel resource of the target BS. If so, it simply forwards the MRHOnotify to the target BS. The BS uses the information of the mobile network to execute fast layer 2 network entry process that eliminates unnecessary information exchange between the BS and the MR (step 5). Once the process completes, all packets whose prefix is the (old) $prefix_{MR}$ are redirected to the mobile network through the target BS. The MNNs can now continue the on-going communications with their CNs (step 6). Afterwards, the MR receives RA from the leaf-FMA (step 7), gets a new $prefix_{MR}$ and broadcasts the $prefix_{MR}$ (step 8 and 9). The MNN then performs the normal Mobile IPv6 operations (steps 10). In the meantime, messages are sent by the SWF for the release of the resources at routers on the previous routing path. At this point, the routing rules inserted by the FMAs for special routing of the mobile network must be removed. The FMAs can start a timer when setting the routing rule and restart the timer when the rule is hit. After an interval when no packets are sent to the old $prefix_{MR}$, which means the MNN has completed the re-registration processes, the routing rules can be removed. In this process, because of the simultaneous execution of finding an optimal new route and performing resource allocation, the handover delay is thus minimized.

In addition, HiMIP-NEMO supports an efficient reactive handover procedure. If the target BS does not receive an MRHOnotify, it sends an MRinfo. FMAs that do not have any information about the MR follow the same procedure as step 2 in Figure 2. Eventually, one of the FMAs will find a record of the MR, and it will become an SWF. The SWF stops the forwarding of newMRquery and preforms the following works: *a)* making QoS reservations for the MR; *b)* adding a routing rule to its routing table for the (old) $pre\text{-}fix_{MR}$; *c)* sending a newMRreply with the newMR field set to “false” and the QoS_reservation_confirm field set to “true” to the leaf-FMA; and *d)* sending an **MRHOinform** message to the root-FMA with the MAC_MR, $prefix_{MR}$ and QoS_{MR}, indicating an abnormal condition. It then creates a newMRreply message and sends along the path back to leaf-FMA. Each intermediate FMA makes corresponding QoS reservation. Once the new QoS route is established, the communications of the MNNs proceed without interruption. In the meantime, messages are sent to release the resources allocated at the routers on the old routing path.

4. PERFORMANCE EVALUATION

We have implemented HiMIP-NEMO on an ns2 simulator that is patched with Mobile IPv6 and IEEE 802.16e modules. A serving BS will send MRHOnotify to the target BS when executing an 802.16e handover, which is suggested but not defined in the IEEE 802.16e standard. To compare the performance, we have implemented Mobile IPv6 Fast Handovers [9] and also Mobile IPv6 Fast Handovers over IEEE 802.16e Networks [17] for NEMO-bs (FNEMO-bs) on ns2. We have implemented a buffer mechanism on the NAR in case that the tunneled packets arrive the NAR before receiving a UNA sent by the MR.

The radius of an IEEE 802.16e cell is 10 kilometers, and the overlap of two adjacent cells is 200 meters. A vehicle with moving speed: V_{MR} km/h first starts up in the coverage of BS1 and then moves to the coverage of BS2. When crossing the boundary of BS1 and BS2, there are three signal strength thresholds for the MR to send MOB_SCN_REQ, MOB_MSHO_REQ and MOB_HO-IND, and they are mapped to three corresponding positions. In other words, the MR will send MOB_SCN_REQ when the distance between BS1 and the MR is 9870 meters, and send MOB_MSHO_REQ and MOB_HO-IND at 9930 and 9930+ α meters, respectively. Note that once the MR receives FBack, it will send MOB_HO-IND immediately; however, if the threshold of sending MOB_HO-IND is reached, the MR will send MOB_HO-IND despite that the FBack is not received.

A two-level FMA hierarchy is used in our simulation, and the link delay between leaf-FMA and the root-FMA is set to 10 ms to simulate multiple hops in real deployment. Leaf-FMA1 and leaf-FMA2 (when using FNEMO-bs, PAR and NAR, respectively) broadcast RAs containing different prefixes through BS1 and BS2 respectively. We assume that in NEMO-bs the home networks of the MR and the MNN are the same. When using NEMO-bs in our simulation, a binding update (BU) will be sent by the MR to the CN which is communicating with the MNN, resulting optimized route. In addition, the delay of duplicate address detection (DAD) at new access router (NAR) is set to 1 second when using FNEMO-bs. A G.711 session is established from a CN to a MNN, which is CBR over UDP with 64 kbps and 20 ms packet interval. Handover delay is defined as the interval between the time that the MNN receives the last packet and the time that the MNN receives the first packet when the MR is attached to BS1 and BS2 respectively. Assume that the IEEE 802.16e handover of the MR is successful.

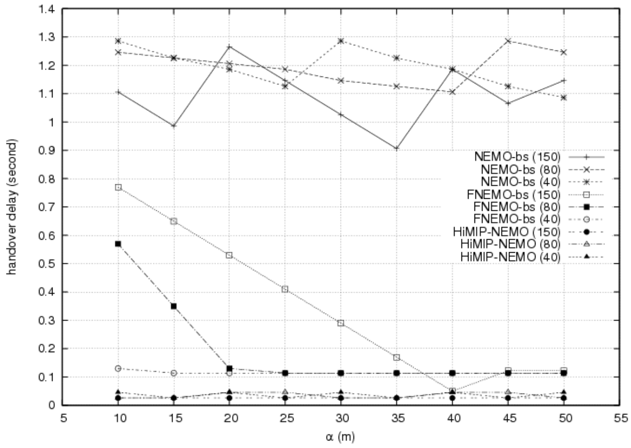

Figure 4. The handover delay.

The handover delays of NEMO-bs, FNEMO-bs and Hi-MIP-NEMO are illustrated in Figure 4. V_{MR} is set to 40, 80 and 150km/h. In NEMO-bs, although the MR is already attached to BS2, it sends a router solicitation after the old RA expires, and then receives a new RA, configures new CoA and sends BU to the HA and the CN. Although the RA interval and the RA lifetime is adjusted to a much smaller value than those defined in Neighbor Discovery in IPv6 [18], the time-consuming processes still lead to a high handover delay. In FNEMO-bs, if α is too small, for example $\alpha \leqq 35$ meters when V_{MR} = 150 km/h, the time interval between reaching the threshold of sending MOB_MSHO_REQ and MOB_HO-IND will too short, and the negotiation between the PAR and the NAR will not be completed in time, which is dominated by DAD on the NAR. Therefore, the MR will execute FNEMO-bs reactive handover which has higher handover delay than FNEMO-bs predictive handover. On the other hand, the HiMIP-NEMO proactive handover is triggered by the layer2 HOind message, and the process time is so short that it can be completed before the MR attaches to BS2. In other word, the handover delay is independent of α. Thus HiMIP-NEMO proactive handover supports high velocity vehicles.

Figure 5 shows the end-to-end packet transmission delay when α is 45 meter and V_{MR} is 150 km/h. There is no packet loss in both FNEMO-bs and HiMIP-NEMO. However, in FNEMO-bs, both of the (tunneled and then) buffered packets in the NAR and the subsequent tunneled packets cause transmission delay fluctuation that leads to high jitter. Furthermore, out-of-order packets are observed after the MR sends BUs to the CN and the HA, because the transmission delay of the packet passing the PAR-NCoA tunnel is

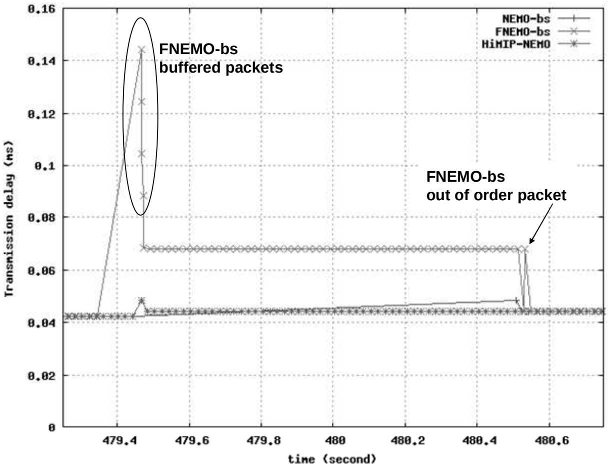

Figure 5. end-to-end transmission delay

higher than that of the packet through optimal path. Both of these two phenomena will degrade the performance of the G.711 session. On the other hand, the routing path remains optimal in HiMIP-NEMO. The communication between the CN and MNN will not be affected when the mobile network moves.

5. CONCLUSION

In this paper, we present a QoS-integrated cross-layer hierarchical network mobility management architecture and the protocols, HiMIP-NEMO, which demonstrates the advantages of combining optimal route-redirecting, resource allocation with network mobility management. Simulation results show that HiMIP-NEMO reduces latency and packet loss for mobile network. Because the network side is responsible for all handover negotiations, including resource reservation, HiMIP-NEMO supports high velocity vehicles, and, in addition, the function of the MR is simplified which benefit network mobility service deployment. Furthermore, an MNN can use the network mobility service without modification, including its security mechanism, e.g. IPsec.

6. REFERENCES

[1] D. Johnson, C. Perkins, and J. Arkko, Mobility support in IPv6, Jun. 2004, IETF RFC 3775.

[2] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thu-bert, Network mobility basic support protocol, Jan. 2005, IETF RFC 3963.

[3] Ernst T., "Network Mobiltiy Support in IPv6" Phd Thesis, University Joseph Fourier Grenoble, France, October 2001.

[4] M. Calderon, C. J. Bernardos, Ma. Bagnulo, I. Soto, and A. de la Oliva, "Design and Experimental Evaluation of a Route Optimization Solution for NEMO," IEEE JSAC, vol. 24, issue 9, 2006

[5] Yu-Chee Tseng, Jen-Jee Chen, and Yu-Li Cheng, "De-sign and Implementation of a SIP-Based Mobile and Vehicular Wireless Network With Push Mechanism," IEEE Transactions on Vehicular technology, vol. 56, NO. 6, NOV. 2007.

[6] Feriel Mimoune, Farid Nait-Abdesselam, Tarik Taleb and Kazuo Hashimoto, "Route Optimization for Large Scale Network Mobility Assisted by BGP," in the IEEE GLOBECOM 2007 proceedings.

[7] C.-M. Huang, C.-H. Lee and J.-R. Zheng, "A Novel SIP-Based Route Optimization for Network Mobility," IEEE Journal on Selected Areas of Communications, vol. 24, issue 9, 2006.

[8] S. Thomson, T. Narten and T. Jinmei, IPv6 Stateless Address Autoconfiguration, Sep. 2007, IETF RFC 4862.

[9] R. Koodli, Ed., Mobile IPv6 Fast Handovers, Jun. 2008, IETF RFC 5268.

[10] R. Braden, L. Zhang, S. Berson, S. Herzog and S. Jamin, "Resource ReServation Protocol (RSVP) version 1 functional specification," RFC 2205 IETF, Sep. 1997.

[11] A. K. Talukdar, B. R. Badrinath and A. Acharya, "MRSVP: a resource reservation protocol for an inte-grated services network with mobile hosts," Wireless Networks 7(1) p5-19, 2001.

[12] C. C. Tseng, G. C. Lee, and R. S. Liu, "HMRSVP: A Hierarchical Mobile RSVP Protocol," Wireless Networks 9, 95-102, 2003.

[13] Tlais, M.; Labiod, H.; "Resource reservation for NEMO networks," International Conference on Wireless Networks, Communications and Mobile Computing, 2005.

[14] Jui-Tang Wang; Yuan-Ying Hsu; Chien-Chao Tseng; "A Band-width-sharing Reservation Scheme to Support QoS for Network Mobility," IEEE International Conference on Communications, 2006.

[15] http://www.ietf.org/html.charters/nsis-charter.html.

[16] IEEE 802.16e, IEEE Standard for Local and Metropoli-tan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Amendment 2: Physical and Medium Access Control Layers for Com-bined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1, 28 Feb. 2006.

[17] H. Jang, J. Jee, Y. Han, S. Park and J. Cha, "Mobile IPv6 Fast Handovers over IEEE 802.16e Networks," Jun. 2008, IETF RFC 5270.

[18] T. Narten, E. Nordmark, W. Simpson and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)," Sep. 2007, IETF RFC 4861.