Handover latency comparison of `SIGMA`, FMIPv6, HMIPv6, and FHMIPv6

Shaojian Fu, **Mohammed Atiquzzaman**
Telecommunications and Networks Research Lab
School of Computer Science, University of Oklahoma,
Norman, OK 73019-6151, USA.
Email: {`sfu,atiq`}`@ou.edu`

Abstract— **In our earlier study, we proposed `SIGMA`, a Seamless IP diversity based Generalized Mobility Architecture. `SIGMA` utilizes IP diversity to achieve a seamless handover of a mobile host, and is designed to solve many of the drawbacks of Mobile IP. In this paper, we compare the handover latency of `SIGMA` and recent MIPv6 enhancements, namely, FMIPv6, HMIPv6, and FHMIPv6. Various parameters are considered such as layer 2 handover/setup latency, IP address resolution latency, layer 2 beacon period, and mobile host moving speed. Our results show that `SIGMA` handover latency is insensitive to layer 2 setup latency, IP address resolution latency and beacon periods. Moreover, `SIGMA` is able to seamlessly handle relatively high speed movement of mobile host.**

I. INTRODUCTION

Mobile IP (base MIP, MIPv6) [1], [2] are the standards proposed by IETF to handle mobility of Internet hosts for mobile data communication. Several drawbacks exist when using MIP in a mobile computing environment, the most important issues of MIP identified to date are high handover latency, and high packet loss rate. Recently, a number of enhancements for MIPv6 are proposed. Fast Handovers for Mobile IPv6 (FMIPv6) [3], aims to reduce the handover latency by configuring new IP addresses before entering the new subnet. Hierarchical MIPv6 mobility management (HMIPv6) [4] introduces a hierarchy of mobile agents to reduce the registration latency and the possibility of an out-dated care-of address. FMIPv6 and HMIPv6 can also be used together as suggested in [4] to improve the performance further (in this paper, we refer to this combination as FHMIPv6). Even with these enhancements, Mobile IP still can not completely solve the high latency problem, and the resulting packet loss rate is still high [5].

We designed a new scheme for supporting low latency, low packet loss mobility called Seamless IP diversity based Generalized Mobility Architecture (`SIGMA`) [6]. It can also cooperate with normal IPv4 or IPv6 infrastructure without the support of Mobile IP. The basic idea of `SIGMA` is to exploit IP diversity to keep the old path alive during the process of setting up the new path to achieve a seamless handover. However, there are some practical obstacle to realizing this principle:

- In the state-of-the-art mobile systems such as IEEE 802.11, GPRS, UMTS, etc. there exists layer 2 handover/setup latency, which is due to the physical and/or link layer limitations. The `SIGMA` signaling messages will experience an extra delay, which may break the parallelism that we hope to achieve with IP diversity.
- After MH move into a new IP domain, it requires some time for MH to obtain a new IP address through DHCP, DHCPv6, or IPv6 Stateless Address Auto-configuration (SAA) [7]. Until this process is finished, MH can not perform any `SIGMA` signaling.
- If MH's moving speed is too high, there is no time for MH to prepare for the new path, the parallelism that can be achieved by IP diversity will be broken.

Therefore, the handover performance of `SIGMA` may affected by these factors mentioned above, even though `SIGMA` does not require any change on the layer 2 or layer 3 implementation. As a comparison, these factors also have impacts on MIPv6 enhancements including FMIPv6, HMIPv6, and FHMIPv6. The *objective* of this paper is to look into the impact of these factors on the handover latency of `SIGMA` and the MIPv6 enhancements. As in paper [6], we illustrate `SIGMA` using SCTP since multihoming is a built-in feature of SCTP.

The rest of this paper is structured as follows: Sec. II outlines the handover signalling procedures of `SIGMA`. The general impact of layer 2 handover latency on `SIGMA` is discussed in Sec. III. The *ns*-2 simulation setup is described in Secs. IV. Sec. V illustrates the impact of layer 2 setup latency on `SIGMA` handover performance through packet trace and congestion window trace. The handover latency comparison of `SIGMA`, FMIPv6, HMIPv6, and FHMIPv6 under various input parameters are shown in Sec. VI. Finally, concluding remarks are presented in Sec. VII.

II. ARCHITECTURE OF `SIGMA`

A typical mobile handover in `SIGMA` using SCTP as an illustration is shown in Fig. 1, where the Mobile Host (MH) is multi-homed node connected through two wireless access networks. Correspondent node (CN) is a single-homed node sending traffic to MH.

The research reported in this paper was funded by NASA Grant NAG3-2922.

0-7803-9415-1/05/$20.00 (C) 2005 IEEE

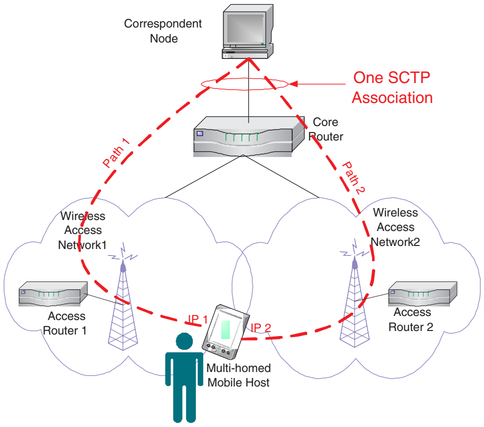

Fig. 1. An SCTP association with multi-homed mobile host.

The handover process of `SIGMA` can be described by the following five steps.

STEP 1: Obtain new IP address

Refer to Fig. 1 as an example, the handover preparation procedure begins when MH moves into the overlapping radio coverage area of two adjacent subnets. Once the MH receives the router advertisement from the new access router (AR2), it should begin to obtain a new IP address (IP2 in Fig. 1).

STEP 2: Add IP addresses into the association

After the MH obtained the IP address IP2 by STEP 1, MH should notify CN about the availability of the new IP address through ASCONF chunks defined by SCTP Address Dynamic Reconfiguration option [8].

STEP 3: Redirect data packets to new IP address

When MH moves further into the coverage area of wireless access network2, CN can redirect data traffic to new IP address IP2 to increase the possibility that data can be delivered successfully to the MH. This task can be accomplished by sending an ASCONF from MH to CN, through which CN set its primary destination address to MH's IP2.

STEP 4: Update location manager (LM)

`SIGMA` supports location management by employing a location manager which maintains a database recording the correspondence between MH's identity and MH's current primary IP address. MH can use any unique information as its identity such as home address like MIP, or domain name, or a public key defined in Public Key Infrastructure (PKI).

STEP 5: Delete or deactivate obsolete IP address

When MH moves out of the coverage of wireless access network1, no *new* or *retransmitted* data should be directed to address IP1. In `SIGMA`, MH notifies CN that IP1 is out of service for data transmission by sending an ASCONF chunk to CN to delete IP1 from CN's available destination IP list.

III. GENERAL ANALYSIS OF FACTORS THAT AFFECT HANDOVER LATENCY ON `SIGMA`

A. *Layer 2 handover/setup concept*

In the state-of-the-art mobile system technologies, when a mobile host changes its point of attachment to the network, it needs to perform a layer 2 (data link layer) handover/setup. As an example, in IEEE802.11 WLAN infrastructure mode, this layer 2 handover will require several steps: detection, probe, and authentication and reassociation with new AP. These procedures can take up to 600-700ms [9] to perform layer 2 handover, after which higher layer protocols can proceed with their signaling procedure. The difference between layer 2 handover and setup is that in setup case the last step is association instead of reassociation in the case of handover. The authors of [9] also show that the most majority of the layer 2 handover time is for detection and channel probing. Therefore, we assume the time required for layer 2 handover and setup are the same. The MIPv6 enhancements have to perform layer 2 handover to cutoff with the old access point and re-associate with a new one since MH has only one interface card, whereas `SIGMA` generally performs layer 2 setup on second interface card while using one card for communicating with old AP.

B. *Impact of layer 2 setup latency, IP address resolution latency, MH moving speed on* `SIGMA`

In `SIGMA`, the layer 2 setup and IP address resolution will postpone the time that MH can start STEP1 (obtain new IP address), since only after layer 2 handover finishes, MH can receive the router advertisement from the new AR. Therefore the STEP2 is also postponed because this step is in synchronous with the STEP1. However, the time of starting STEP3 and STEP4 may or may not be affected by the layer 2 handover latency. Consider a linear movement from AR1 to AR2 as an example, ideally (without any layer 2 handover latency and IP address resolution latency) the STEP3 and STEP4 of `SIGMA` handover should start at (say time t) the point of the overlapping region that gives MH enough time to finish STEP3 and STEP4 before it moves out of the coverage of AR1. When layer 2 setup latency and IP address resolution latency come into play, depending on the MH's moving speed, overlapping region size, round trip time from MH to CN (for ADDIP chunks to come back), the time (say time t') that STEP2 finishes could fall before or behind the time t. If $t' \leq t$, the layer 2 setup latency and IP address resolution latency has virtually no impact on `SIGMA` handover since the new data path through AR2 is available before MH moves into coverage of AR2, and there is no loss happened due to `SIGMA` handover. However, if $t' > t$, the layer 2 setup push the latest starting point of STEP3 and STEP4 from t to t', which will cause these two steps cannot be finished before MH moves out of AR1 coverage, and some packet losses will happen.

IV. SIMULATION TOPOLOGY AND PARAMETERS

In this section, we describe the simulation topology and parameters that have been used to compare the performance of `SIGMA`, FMIPv6, HMIPv6, and FHMIPv6. We have used *ns-2* simulator that supports SCTP as the transport protocol. We implemented `SIGMA` protocol for *ns-2*, and incorporated FMIPv6, HMIPv6, FHMIPv6 implementations used by [5] to support the simulation comparison. Standard *ns-2* simulator does not have direct support for layer 2 handover latency simulation; an MH can communicate with two APs simultaneously once the MH entering into the overlapping region of the two APs. In order to simulate mobile handovers between real-world infrastructure mode WLANs, we also implemented layer 2 handover latency in *ns-2* IEEE 802.11 code by introducing layer 2 beacons and a set of timers.

A. Simulation topology

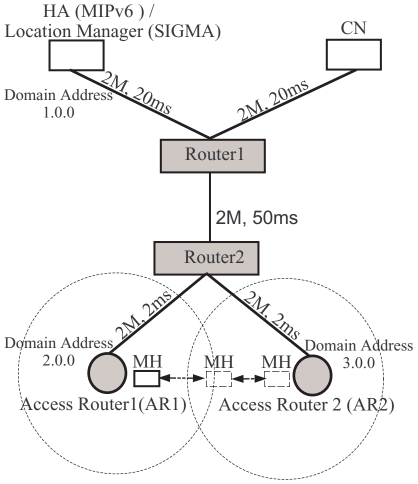

Fig. 2. Simulation topology.

The network topology used in our simulations is shown in Fig. 2. In the figure, AR1 and AR2 stand for two access routers. MH initially has an IP address of 2.0.1 (IP1) when it is associated with AR1. After moving into the overlapping region, MH will get new IP address 3.0.1 (IP2) from AR2, which will make it have two IP (IP1 and IP2) available at the same time. Once MH moves out of the coverage of AR1, the IP1 is deleted and only IP2 is available. In the figure, MIPv6 uses HA, while `SIGMA` uses it as Location Manager. Router2 in the topology will act as an MAP point in HMIPv6 and FHMIPv6, while act as only a normal router in FMIPv6 and `SIGMA`. The link characteristics, namely the bandwidth (Megabits/s) and propagation delay (milliseconds), are shown on the links.

B. Simulation parameters

We have used the following parameters in our simulations:

- A pair of FTP source and sink agents are attached to the CN and MH, respectively, to transfer bulk data from CN to MH. To stabilize the result, each simulation run lasts for 500 seconds of MH's linear back and forth movement between AR1 and AR2.
- Each base station has a radio coverage area of approximately 40 meters in radius. The overlapping region between two ARs is 10 meters.
- To make a fair comparison, we have used standard SCTP protocol (without mobility related modifications) as the transport layer protocol for MIPv6 enhancements. This is to ensure that all the handover schemes use the same connection setup and congestion control control mechanisms, and that the results are only affected by the different handover schemes.

V. EFFECT OF LAYER 2 SETUP LATENCY ON `SIGMA`

In this section, we will show simulation packet traces and congestion window traces of `SIGMA` to illustrate the impact of layer 2 setup latency on `SIGMA` handover performance. These trace results can be classified into three categories: (1) no layer 2 setup latency, (2) layer 2 setup latency does not cause packet loss in `SIGMA` handover, (3) layer 2 setup latency introduce some packet losses in `SIGMA` handover. In all categories, the IP address resolution latency is set to a large value of 500ms to cover the scenarios where getting IP address may take long time.

A. No layer 2 setup latency

Fig. 3 shows the packet trace observed at the CN during one typical handover for `SIGMA` with data being sent from CN to MH. Layer 2 setup has no latency, i.e. it finishes immediately. The segment sequence numbers are shown as MOD 100. From Fig. 3 we can observe that SCTP data segments are sent to MH's IP1 until time 8.140 sec (point t_1), then the IP2 almost immediately (point t_2), and all these packets are successfully delivered to MH. Since the change of routing table at MH takes at the same time as the sending of SetPrimary chunk to CN at STEP3 in Sec. II, the ACKs sent to CN after time 8.134 sec (the time handover decision is made) use the new path through AR2, which is not the same as the path receiving the data packets before time 8.140 sec. Also note that at t_2 a slow start begins at address IP2. The initial congestion window ($cwnd$) is three instead of two (as specified in RFC2960) because CN has received an ACK from the new path and $cwnd$ is increased by one segment size. The next window of data is sent to IP2 at time 8.40 sec using $cwnd$ of six according to slow start algorithm.

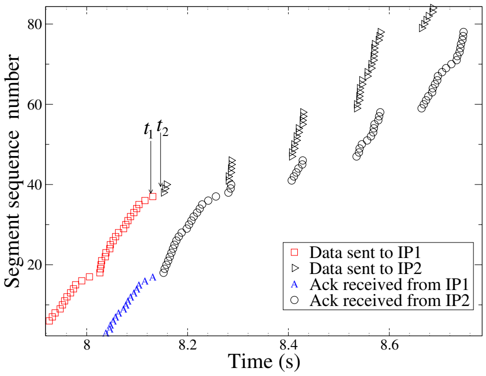

Fig. 3. Segment sequence of `SIGMA` during one handover with no layer 2 setup latency.

B. Low layer 2 setup latency

Fig. 4 shows the packet trace observed at the CN during one typical handover for `SIGMA` with layer 2 setup latency of 200ms. From Fig. 4 we can observe that SCTP data segments are sent to MH's IP1 until time 8.16 sec (point t_1), then the IP2 almost immediately (point t_2), and all these packets are successfully delivered to MH. Therefore, `SIGMA` still experienced a seamless handover because it can prepare the new path in parallel with data forwarding over the old path. We found that in this kind of scenario *the only impact of layer 2 setup latency is to push the time instant of transport layer handover by 20ms* (8.14 sec vs. 8.16 sec). This is the basic reason that explains why `SIGMA` can achieve a low handover latency, low packet loss rate and high throughput as shown in [6].

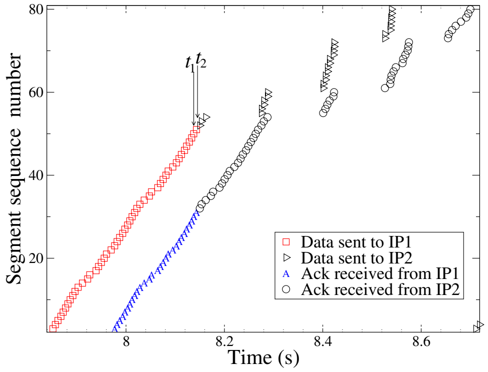

Fig. 4. Segment sequence of `SIGMA` during one handover with layer 2 setup latency of 200ms.

C. High layer 2 setup latency

Fig. 5 shows the packet trace observed at the CN during one typical handover for `SIGMA` with layer 2 setup latency of 500ms. From Fig. 5 we can observe that all SCTP segments sent to address IP1 starting at t_1 until the end of the window are all lost. The reason for this is that layer 2 setup postpone the preparation of new path, while the old path becomes unavailable after time 9 sec. The RTO value for the old path at this time is 1.0 sec. Therefore, at time t_2 (around time 10.0 sec.), the first lost segment is retransmitted to the new path, which is delivered successfully. However, the `SIGMA` handover still have not finished by this time, and the routing table from MH to CN still requires the ACK go through the old path, which is lost again. This will make the RTO of the new path doubled to 2.0 sec. The next retransmission happens at the old path. This time the initial RTO value of new path will be used: 3.0 seconds as specified by RFC2960, which results in the retransmission taking place at time 13.0 sec (10.0+ RTO value of 3.0 of new path). This retransmitted packet is also lost since the old path is not available at that time. Only after time 15 sec. (13.0+RTO value 2.0 at old path) the third retransmission make the association back to the normal transmission.

VI. HANDOVER LATENCY COMPARISON RESULTS

In this section, we present comparison results showing the effect of various input parameters on the handover latency of `SIGMA` and compare with MIPv6 enhancements. We define *handover latency* as the time interval between the last data segment received through the old path and the first data segment received through the new path from CN to MH.

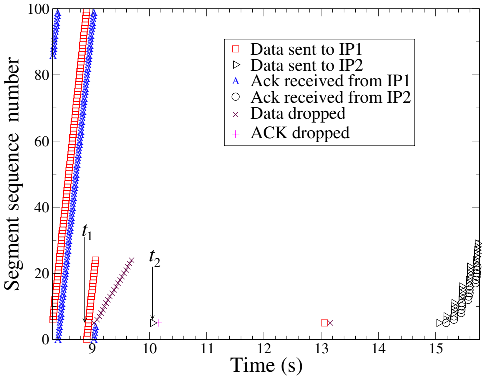

Fig. 5. Segment sequence of `SIGMA` during one handover with layer 2 setup latency of 500ms.

A. Impact of L2 handover/setup latency and address resolution latency

First we look at the overall handover latency of `SIGMA` compared with MIPv6 enhancements when the layer 2 handover/setup latency range from 100 to 600ms, and IP address resolution latency ranges from 300 to 600ms, as shown in Fig. 6. The IP address resolution latency is denoted as α in the figure. The values of layer 2 handover/setup latency corresponds to the empirical values in IEEE 802.11 networks [9]. The moving speed is fixed at 5m/s.

It can be seen from Fig. 6 that the handover latency of `SIGMA` is very low (in the range of 5-10ms) when the combined latency of layer 2 setup and IP address resolution is less than 900ms. This is because when the MH is using the old path to do communication with CN, it can perform the layer 2 setup and IP address resolution on the other interface in parallel (as shown in packet trace in Sec. V-A and V-B), thus the impact of these latencies can be noticeably reduced compared to enhancements of MIPv6. When the combined latency is larger than 900ms, this parallelism is broken since the MH does not have enough time to finish all the signaling required in `SIGMA`. Some packets sent to the outdated AR are lost and CN is forced to backoff by SCTP's congestion control algorithms. The packet trace in Sec. V-C shows the example where high layer 2 latency causes packet losses and high `SIGMA` handover latency.

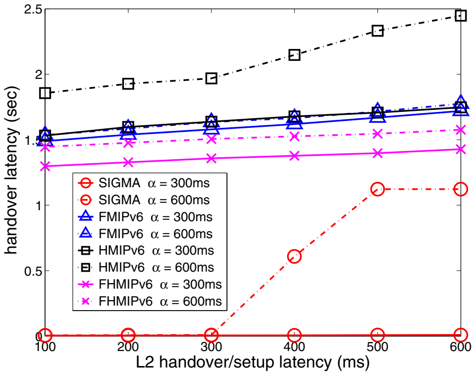

Fig. 6. Impact of layer 2 handover/setup latency and IP address resolution latency

It can be also observed from Fig. 6 that the handover latency of MIPv6 enhancements is around 1.40-2.49 seconds, which is much higher that that of SIGMA. This is because even FMIPv6 and FHMIPv6 can perform address resolution and prepare tunnelling between two ARs in advance, MH still cannot receive packets from the new path before completion of layer 2 handover. The resulting packet loss will force the CN to backoff and postpone the time that MH can receive the packet from the new path.

For FMIPv6 and FHMIPv6, MH can perform the address resolution in advance, which will reduce the impact of address resolution latency on the overall handover latency. For HMIPv6, neither layer 2 handover latency nor IP address resolution latency can be avoided. Therefore, when layer 2 handover latency and address resolution latency increase, the overall handover latency for HMIPv6 will increase. Compared with FMIPv6 and FHMIPv6, HMIPv6 is more sensitive to IP address resolution latency.

B. Impact of moving speed and layer 2 beacon period

Next we vary the movement speed of MH from 2.5m/s up to 20m/s, vary the layer 2 beacon period from 20ms to 80ms, and fix both of the layer 2 handover/setup latency, IP address resolution latency to 100ms. Fig. 7 shows the impact of MH's moving speed and layer 2 beacon period (τ) on the overall handover latency of SIGMA, FMIPv6, HMIPv6, and FHMIPv6. When MH's moving speed is less than 15m/s, the impact of moving speed is not obvious for SIGMA. When MH moves faster, SIGMA will experience a higher handover latency due to MH having insufficient time to prepare for the handover. Therefore, there is a higher possibility that the packets are forwarded to the outdated path and get lost, and the time instant that MH can receive the packets from new path will be postponed and the handover latency increases accordingly.

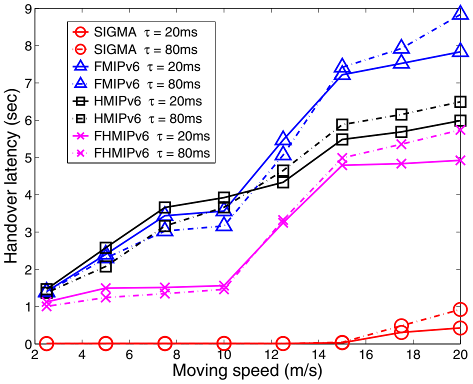

Fig. 7. Impact of moving speed and layer 2 beacons

We can also observe in Fig. 7 that when MH's moving speed is higher, all MIPv6 enhancements will experience a higher handover latency due to less time to prepare for the handover. However, the increase in speed has most significant effect on FMIPv6 and FHMIPv6 since they rely on the assumption that detection of the new agent is well in advance of the actual handover. When the moving speed is higher, the assumption can break down more easily. Because HMIPv6 and SIGMA do not rely on this assumption, the effect of moving speed is smaller. But when moving speed is higher, there is higher possibility that the packets are forwarded to the outdated path and get lost, therefore the time instant that MH can receive the packets from new path will be postponed and the handover latency increases accordingly.

Comparing the curves of different layer 2 beacon period in Fig.7, we can see a layer 2 beacon period of 20ms generates the highest handover latency at low moving speeds (under 10m/s). This is because too low a beacon period (e.g. 20ms) produces a high volume of beacons, which will contend for the limited wireless bandwidth with data and signaling traffic. The packet loss rate for the signaling packets thus increase and it may require additional retransmission time to deliver them successfully. The resulted handover latency will therefore be increased. However, at higher speed (more than 15m/s), the low layer 2 beacon period can help the MH to detect the new AP and begin layer 2 handover/setup earlier, thus reduce the possibility that packets are forwarded to outdated path. The resulted handover latency decreases accordingly.

VII. CONCLUSIONS

This paper compares the handover latency of SIGMA and MIPv6 enhancements through simulation. The impact of different input parameters, including layer 2 handover/setup latency, IP address resolution latency, MH moving speed, and layer 2 beacon period, have been investigated. Our results indicate that for typical network configuration and parameters, SIGMA is not sensitive to layer 2 setup latency, IP address resolution latency and layer 2 beacon periods. The handover latency of SIGMA is lower than that of MIPv6 enhancements under all of the simulated scenarios. SIGMA has also been shown to be able to seamlessly handle relatively high speed movement.

REFERENCES

[1] C.E. Perkins, "Mobile Networking Through Mobile IP," *IEEE Internet Computing*, vol. 2, no. 1, pp. 58–69, January/February 1998.
[2] D. Johnson, C.E. Perkins, and J. Arkko, "Mobility support in IPv6." IETF RFC 3775, June 2004.
[3] R. Koodli (editor), "Fast handovers for Mobile IPv6." IETF DRAFT, draft-ietf-mipshop-fast-mipv6-03.txt, October 2004.
[4] H. Soliman, C. Catelluccia, and K.E. Malki et al., "Hierarchical Mobile IPv6 mobility management (HMIPv6)." IETF DRAFT, draft-ietf-mipshop-hmipv6-04.txt, December 2004.
[5] R. Hsieh and A. Seneviratne, "A comparison of mechanisms for improving Mobile IP handoff latency for end-to-end TCP," *ACM MobiCom*, San Diego, USA, pp. 29–41, September 2003.
[6] S. Fu, L. Ma, M. Atiquzzaman, and Y. Lee, "Architecture and performance of SIGMA: A seamless mobility architecture for data networks," *40th IEEE Internationl Conference on Communications (ICC)*, Seoul, Korea, May 2005.
[7] S. Thomson and T. Narten, "IPv6 stateless address autoconfiguration." IETF RFC 2462, December 1998.
[8] R. Stewart, M. Ramalho, and Q. Xie et. al., "Stream control transmission protocol (SCTP) dynamic address reconfiguration." IETF DRAFT, draft-ietf-tsvwg-addip-sctp-09.txt, June 2004.
[9] A. Mishra, M. Shin, and W. Arbaugh, "An empirical analysis of the IEEE 802.11 MAC layer handoff process," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 2, pp. 93–102, April 2003.