Hindawi Publishing Corporation EURASIP Journal on Wireless Communications and Networking Volume 2008, Article ID 348594, 8 pages doi:10.1155/2008/348594

Research Article

A Two-Layered Mobility Architecture Using Fast Mobile IPv6 and Session Initiation Protocol

Deeya S. Nursimloo, George K. Kalebaila, and H. Anthony Chan

Department of Electrical Engineering, University of Cape Town, Rondebosch 7701, South Africa

Correspondence should be addressed to Deeya S. Nursimloo, deeya31@gmail.com

Received 27 November 2006; Accepted 24 May 2007

Recommended by Kameswara Rao Namuduri

This paper proposes an integrated mobility scheme that combines the procedures of fast handover for Mobile IPv6 (FMIPv6) and session initiation protocol (SIP) mobility for realtime communications. This integrated approach is based on the context of the applications utilized. Furthermore, to reduce system redundancies and signaling loads, several functionalities of FMIPv6 and SIP have been integrated to optimize the integrated mobility scheme. The proposed scheme aims at reducing the handover latency and packet loss for an ongoing realtime traffic. Using ns-2 simulation, we analyze the performance of the proposed integrated scheme and compare it with the existing protocols for a VoIP and for a video stream traffic. This mobility architecture achieves lower handover delay and less packet loss than using either FMIPv6 or SIP and hence presents a powerful handover mobility scheme for next generation IP-based wireless systems.

Copyright © 2008 Deeya S. Nursimloo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

The next-generation wireless systems are envisioned to have an IP-based infrastructure platform to support the heterogeneity of the access technologies. Currently, various wireless technologies and networks provide different services to mobile users based on their requirements.

Mobility management is required to enable seamless roaming among the heterogeneous networks and to minimize service disruptions in the realtime applications during handover. In a heterogeneous environment, mobilityenabled protocols are considered to achieve global roaming among the various access technologies. Currently, the two leading approaches to support mobility of services in the IP core network are Mobile IP (MIP), which supports mobility across the network layer, and session initiation protocol (SIP), which supports mobility through the application layer. Yet both protocols suffer from different types of drawbacks that impact on the media flow during the handover mechanism. Fast handovers for Mobile IPv6 protocol is one of the proposed enhancements of Mobile IP within the IETF Mobile IP working group. Its performance is based on the capability of supporting two types of handover: reactive and proactive handover mechanisms. The proactive

mechanism aims to reduce service degradation that a mobile device could suffer due to a change in its point of attachment. SIP is an application layer protocol, which allows the provisioning of services in IP-based networks. Therefore, there is a need to seamlessly interwork fast Mobile IP and SIP to support mobility transparency to realtime services.

This paper proposes an integrated mobility scheme that combines procedures of fast handover for Mobile IPv6 and SIP mobility for realtime communications. An analysis of the protocols is presented to support terminal mobility. Furthermore, to reduce the system redundancies and signaling load, several functionalities of fast Mobile IP and SIP have been integrated to optimize the mobility architecture. The rest of the paper is organized as follows: Section 2 surveys the related background work on IP mobility protocols. The proposed mobility design is outlined in Section 3. Section 4 describes the simulation results generated from ns-2. Finally, Section 5 summarizes and concludes the paper.

2. IP MOBILITY

This section describes the previous work related to fast Mobile IPv6 and SIP.

2.1. Fast handover for Mobile IPv6

Fast handovers for Mobile IP protocol proposed by the Mobile IP working group of the IETF [1] specify the enhancements to Mobile IPv6 that enable a mobile node (MN) to connect to a new point of attachment more rapidly. The protocol aims to reduce service degradation by minimizing the time during which an MN is unable to send or receive IP packets. With the emergence of realtime traffic, it is necessary to ensure IP connectivity and rapid handovers to avoid unnecessary latencies. In the proactive mechanism, the mobile node acquires information about its new access router prior to moving to it. When the mobile node is detected by the new access router, a new link is already established to send and receive data packets.

2.2. Predictive handover—FMIPv6

A predictive-determined handover is when the MN is responsible for defining and initiating the handover prior to the handover as illustrated in Figure 1.

To initiate the fast handover mechanism, the MN sends an RtSolPr message to the previous access router (PAR) to indicate that a handover is required to move to its next point of attachment. The RtSolPr message contains the link layer address or the identifier of its new point of attachment. The PAR will reply with a *PrRtAdv* message which informs the MN of the new care-of address (CoA) that will be used to deliver the packets together with the IP address and link layer address of the new access router (NAR). In addition to the above message, the PAR sends a HI message to the NAR with both the new configured CoA and the old CoA that was used at the PAR. The NAR checks whether the newly formulated CoA is a valid address to ensure that it has no duplicate. If the new CoA is valid, the NAR adds it to the neighboring cache entry and responds with a HAck message. The MN sends a fast binding update (FBU) to the PAR to confirm that the handover is to take place. On receipt of the FBU and of the HAck message, the PAR can initiate the forwarding of the packets destined to the MN's old CoA to either the newly assigned CoA or the NAR. The MN does not use the newly assigned CoA until the fast binding acknowledgement (F-BACK) message is sent through a temporary tunnel. As soon as the MN gains connectivity with the NAR, a fast neighbor advertisement (FNA) message will be sent. This message is to trigger the forwarding of the packets for the MN, assuming that the NAR is aware of the MN or else packets are likely to be dropped. The FNA message contains the old and new CoAs as well as the link layer address. The NAR will check the link layer address to check if there is a mapping in the neighbor cache. The exchange of information between the routers is to facilitate the forwarding of packets and to minimize the latency perceived by the MN during handover.

2.3. Session initiation protocol

SIP is a protocol developed in the IETF by the multiparty multimedia session control (MMUSIC) for establishing multimedia session. SIP is a text-based protocol whose main entities are user agents, proxy servers, redirect servers, and registrars [2]. Call address is defined, for example, by user@host where "user" is the user name and "host" is a domain name. As discussed in [3], SIP supports terminal mobility to establish connection when a mobile node has already moved to a different location or during the middle of a session [4]. Midcall mobility, as shown in Figure 2, is when a mobile node moves during an ongoing session. The terminal will detect a network address change (this is achieved through a DHCP server or a variant of it) and will send a new INVITE message (Re-INVITE) with updated session description protocol (SDP) to the correspondent node without going through intermediate SIP proxies. The INVITE request will inform the remote user of the change in the session parameters with the new IP address to forward the packets correctly. The significant drawbacks on the SIP-based mobility mechanism are the disruptions caused during call setup and the absence of mobility management support for long-term TCP connections.

3. PROPOSED ARCHITECTURE

To provide a complete mobility management framework for realtime applications, it is necessary to combine both network layer protocol FMIPv6 and application layer protocol SIP, in a way to complement each protocol feature based on their kind of application. This section focuses on merging the traditional protocol schemes to form an integrated protocol scheme to support the handover procedures in overlapping networks. We outline the steps involved in providing mobility support in the proposed scheme.

3.1. IP-based handover

The architectural design of the proposed mobility framework aims to provide IP-based handoff management from network layer fast Mobile IP and SIP at the application layer. In that way, it will allow intrinsic connections between low-level and high-level mobility.

The aim of fast handovers for Mobile IPv6 protocol specification is to enable the mobile node to configure a new care-of address before it moves to a new access router. In that way, the new care-of address will allow immediate connection to the new access router, with minimal interruption to the packets flow between the routers. The mobile node will acquire a care-of address in a way that a duplicate or invalid packet address is not picked.

3.2. Address configuration

After completing the layer 2 handover, address configuration may either follow stateful, that is, through DHCP or stateless address reconfiguration procedures [5]. In any case, duplication address detection (DAD) is needed to verify the uniqueness of the address, and this process brings additional delays to the whole handover procedure [6].

Deeya S. Nursimloo et al.

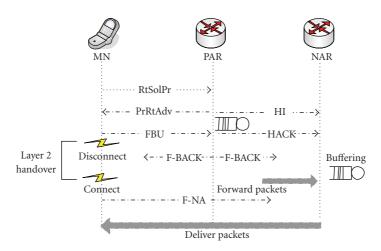


FIGURE 1: Predictive handover mechanism in FMIPv6.

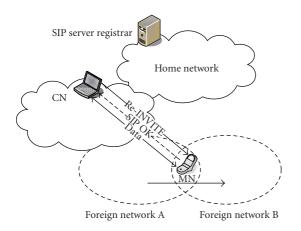


FIGURE 2: SIP-based Mid-call mobility.

3.3. Registration within home agent and SIP registrar

The purpose of the registration in Mobile IP is to inform the mobile node's home agent to register its new care-of address through a binding update (BU) message. In this way, it can also inform the corresponding node (CN) about the new care-of address to appropriately forward/send the packets destined to the mobile node. In the case of TCP or non-SIP applications, the connections can be maintained without a disruption.

An extension of the home agent specification is proposed in the design model in order to colocate the mechanism of the SIP registrar. For the purpose of this research, it is necessary that during an SIP re-establishment session, the correspondent node is informed of the MN's new IP address so that it can communicate directly to the MN. In order to do so, a binding mechanism between the temporary IP address of the MN and the user level identifier is required to update the current location of the MN. Once the current location is updated, the SIP proxy and SIP redirect server database can be updated. The domain name system (DNS) records and

helps in finding SIP proxies responsible for routing the SIP messages to the destination domain.

The home agent (HA) will update the IP address and the user SIP ID to inform the CN of the current location. In the case of a TCP packet being sent through, the home agent will update the CoA and, through route optimization, register to the correspondent node. However, if it is a realtime application transfer, the home agent will update the SIP registrar server with the new location of the mobile node. Upon completion of the handover mechanism using this approach, the HA does not tunnel data for the MN as packets are delivered directly through an RTP connection setup from the CN to the MN.

3.4. SIP session re-establishment

After acquiring a new IP address before handover, the mobile node, as an SIP user client, initiates the handover procedure by sending a Re-INVITE message to the correspondent node. The SIP Re-INVITE message initiates the registration within the SIP registrar at the home network of the mobile node and carries the updated SDP parameters to the CN. As a result, call parameters are renegotiated on an end-to-end basis with the SIP proxy server and SIP redirect server as an intermediate to support soft handover. In this scheme, end-to-end negotiation protocol [7] is implemented within the SIP proxy together with SDPng (SDP extensions) for quality of service coordination. Adaptation will be translated when a change in quality of service (QoS) occurs. The session re-establishment allows the CN to redirect all its ongoing media streams and signaling sessions directly to the MN's current IP address as it attaches to the new point of attachment. The Re-INVITE message, similar to the INVITE message configuration, contains the new IP address and the updated contact field where the MN will receive SIP messages in future. If the correspondent node responds with an SIP OK message, agreeing to IN-VITE response, the MN will in turn respond with an ACK to complete the SIP messaging before data transfer. For realtime applications, it is necessary to decrease delays and packet loss as much as possible, and the integrated scheme aims at

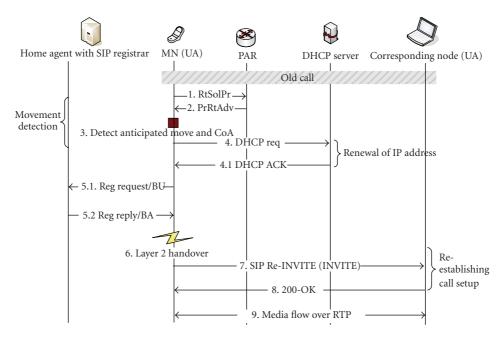


FIGURE 3: Handover signalling flowing using FMIPv6 + SIP.

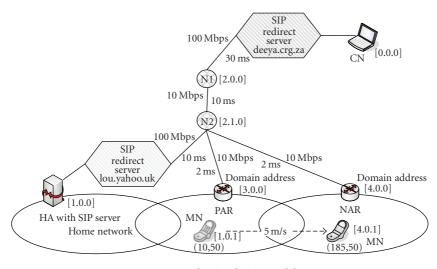


FIGURE 4: The simulation model.

avoiding triangular routing and any kind of encapsulation mechanism during the ongoing calls.

The proposed architecture has both FMIPv6 and SIP mobility procedures simultaneously to provide an integrated handover mobility scheme as explained in the message flow (see Figure 3). The design, as discussed above, aims at reducing the signaling loads by integrating the redundant messages from both protocols for complete message registration for ongoing calls. The transfer of the media flow is accomplished through SIP procedures.

4. SIMULATION TOPOLOGY AND PARAMETERS

In this section, the simulation setup is presented to investigate packet loss, handover latency, and signaling latencies.

The simulations are run using ns-2 version 2.27 [8]. The ns-2 evaluation framework is modified to support the fast Mobile handovers [9] and SIP signaling messages based on the NIST SIP module [10]. Realtime traffic, that is, a VoIP application and streaming of video packets, is characterized to illustrate and compare the performance of the proposed architecture to the existing schemes.

4.1. Simulation model

All the simulations are performed using the network topology as shown in the network simulation topology (see Figure 4).

The following simulation environment consists of a correspondent node (CN), streaming realtime traffic (i.e., VoIP

Deeya S. Nursimloo et al. 5

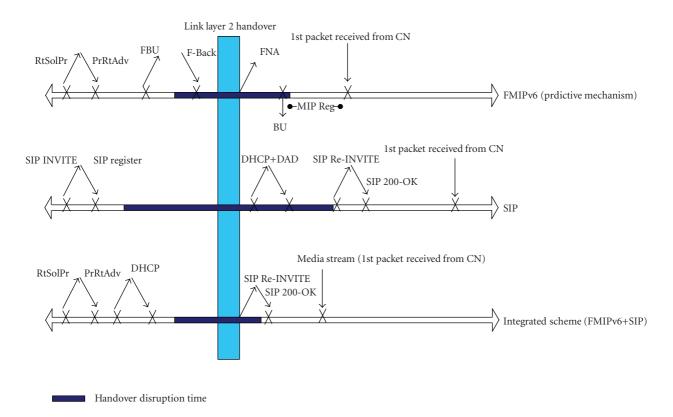


FIGURE 5: Handover signalling flow for FMIP, SIP, and FMIPv6 + SIP.

and video packets) with RTP over UDP setup to a mobile node (MN), home agent (HA) with a colocated registrar and SIP redirect servers. In the case of a small-scale simulation environment, it is not necessary to include a DNS. The SIP redirect server is connected to the CN with a given URL (deeya.crg.za) and the SIP redirect server is connected to the MN with a given URL (lou.yahoo.uk).

The CN is a constant bitrate (CBR) source, transmitting packets in an RTP over UDP medium. The MN acts as a sink, by receiving the packets from the CN at a constant interarrival rate.

A one-way VoIP connection can be modeled by a stream of packets with a fixed packet size and transmission rate. The CN produces packets with a fixed length of 200 bytes made up of a payload of 160 bytes and headers (RTP + UDP + IP) of 40 bytes. A typical PCM voice-coding scheme G.711 is emulated with a packet data rate of 64 kbps corresponding to 20-millisecond frames. The bandwidth and link delay between the two intermediate wired nodes (N1, N2) and the access routers (PAR, NAR) are configured to 10 Mbps and 10 milliseconds, respectively. Between the access routers and the mobile node, these parameters are set to 10 Mbps and 2 milliseconds. From the wired nodes to HA and to CN, the bandwidths are both set to 100 Mbps whereas the link delays are, respectively, set to 10 milliseconds and 30 milliseconds as shown in the simulation model. The movement model for the simulation scenario allows the MN to move linearly between the two access networks. The MN starts to move towards the NAR from PAR at 10 seconds from simulation time, at a speed of 5 m/s.

5. SIMULATION RESULTS

The results of the performance of the proposed integrated scheme are presented and compared to the existing protocols' architectures: FMIP and SIP. The main motivation for the optimization of the proposed scheme is to reduce the delays incurred by the existing protocols during handover.

Figure 5 illustrates the handover signaling disruption timeline of the protocols discussed in the experimental setup. The handover disruption times in FMIPv6 and in the integrated scheme depend largely on the availability of the handover-related information from lower layers to the IP layer. In pure SIP setup, the disruption time is higher because it has no mechanism to indicate the eminent handover. The handover disruption time for FMIP and the integrated mobility framework does not differ much because both use the same handover detection mechanism to indicate eminent handover. The timeline only shows important messages exchanged between the MN and the AR, HA, and SIP agents.

5.1. Handover latency and packet loss

In terms of packet loss as shown in Table 1, the integrated model shows a 37% decrease in packet loss compared to the FMIPv6 predictive mechanism. The integrated model shows

	_	_	
	FMIPv6 (predictive mechanism)	SIP (terminal mobility mechanism)	Integrated mobility model FMIPv6 + SIP
Packet loss	16	13	6
Average handover latency (ms)	100.96	_	110.75
Average throughput (kBytes/s)	63.92	58.54	61.98

Table 1: Comparisons of different protocols schemes.

an improved performance because SIP takes over the reestablishment of media flow after FMIP movement detection mechanism. The significant high packet loss in FMIPv6 as compared to the integrated scheme could be attributed to the time ambiguity problem [11] in FMIP implementation in ns-2. FMIPv6 mechanism performs IP care-of address configuration and prepares for the tunneling before the handover between the two ARs. During that time, the MN cannot receive any packets from the new router before the layer 2 handover takes place. The integrated model experiences a higher handover latency than FMIP, even though it uses the same prediction mechanism as FMIP, because it uses SIP immediately after layer 2 handover to re-establish data flow which takes longer to converge. Although the integrated scheme had higher handover latency, it experienced a 37% decrease in packet loss. This disparity can only be attributed to the implementation of FMIP in the simulation. A much-refined implementation would have resulted in lower packet loss than the integrated scheme. In terms of average throughput, the protocol mechanisms achieve approximately the same system performance after handover.

5.2. Handover signaling latency

Each data point on all graphs shown below corresponds to an average of 20 independent handover simulation events. The handover-associated signaling latency is measured against the distance traveled by the MN with reference to the CN. The graph does not include the delay incurred during DAD execution. Figure 6 illustrates the signaling delay comparison of the proposed integrated scheme to SIP.

The integrated scheme (FMIP+SIP) shows a marked improvement in performance in terms of handover signaling latency compared to SIP. The integrated scheme shows a 42% reduction in handover delay over SIP. This improvement is attributed to the FMIP handover detection mechanism. In the integrated scheme, FMIP is used for movement detection and SIP is used to re-establish the session between the MN and the CN. The CoA is configured before L2 handover enabling the MN to send an SIP Re-INVITE message to the CN immediately after L2 handover is complete. In comparison, the SIP scheme has to wait until L2 handover is complete before it can get the CoA from the DHCP server and then send

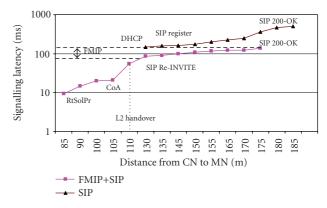


FIGURE 6: Handover signaling delay for VoIP in SIP and FMIP+SIP.

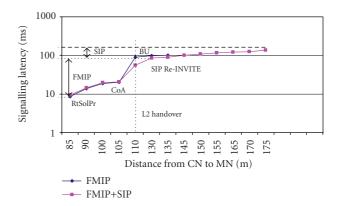


FIGURE 7: Handover signaling delay for VoIP in MIP and FMIP + SIP.

an SIP Re-INVITE message to the CN. The handover signaling in the integrated schemes converges faster than in the SIP scheme owing to the absence of movement detection in the latter. Therefore, the 42% performance improvement in handover delay is attributed to FMIP as shown on Figure 5. All the signalings after L2 handover are SIP signaling messages to re-establish media flow and all signaling before L2 handover are FMIP-related message for movement detection. The 42% handover delay improvement also accounts for the low packet loss the MN experiences during handover as compared to SIP (see Table 1). The main signaling messages ex-4 changed between theMNand CNare as labeled in Figure 3.

FMIPv6 simulation model characterizes VoIP application at a constant bitrate (CBR) with UDP and the integrated scheme supports realtime communication with RTP over UDP. At 8.6 seconds, movement detection mechanism in FMIP is triggered resulting in the MN sending the router solicitation message (RtSolPr). This initiates the FMIP associated signaling to prepare for eminent L2 handover. The MN is then assigned the CoA before L2 handover. In the FMIP scheme, the MN continues with registration with the HA and the CN by sending BUs, whereas in the integrated scheme, the MN waits for L2 handover to complete before re-establishing the media flow through SIP. From Figure 7, there is a marginal difference in performance in terms of handover signaling delay because both schemes use the same

Deeya S. Nursimloo et al.

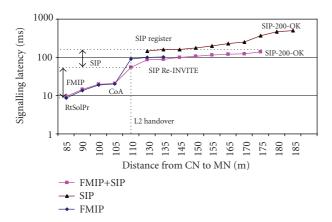


FIGURE 8: Handover signaling delay for FMIP + SIP, SIP, and FMIP.

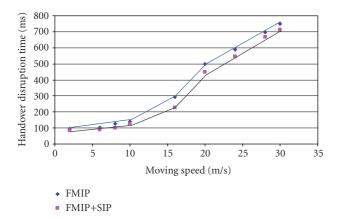


Figure 9: Moving speed of mobile node versus handover latency.

movement detection mechanism. FMIP scheme converges faster than the integrated scheme, which has to go through SIP signaling to re-establish media flow. Therefore, FMIP reestablishes packet flow faster than in the integrated scheme even though we cannot account for the high number of packet loss in FMIP.

Figure 8 combines results from Figures 6 and 7. The proposed integrated mobility model shows an overall reduction in handover signaling latency compared to pure SIP schemes for any type realtime traffic investigated.

5.3. Movement speed

This section investigates the influence of movement speed of the MN on handover disruption time. The MN's speed is varied from 2 m/s up to 30 m/s. From Figure 9, both FMIP and the proposed integrated scheme (FMIP+SIP) are severely affected by the increase in speed although the proposed scheme shows marginal improvement in performance. The result can be attributed to the fact that both schemes employ the same handoff detection mechanism to well detect the new access router in advance of the actual handover.

With increasing movement speed of the MN, the detection time is reduced and thus preparation for the anticipated

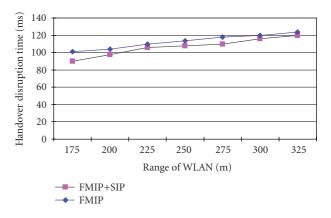


FIGURE 10: Range of WLAN versus handover latency.

handover process cannot be completed in time of the handoff. Though the disruption time is a function of the handoff detection mechanism used, handoff preparation time is protocol-dependent and remains constant as long as the same protocol is used, which in this case is FMIP. This dependence explains the marginal difference in performance of the two schemes. Movement speed also affects packet loss due to handoff. The increase in MN's speed increases the possibility of packets being forwarded to the outdated path and thus increasing the probability of packet loss.

5.4. WLAN range

Figure 10 shows the effect of the different WLAN ranges between the PAR and the NAR on the handover disruption time

Figure 10 shows that range of access points (APs) have only a little effect on the average handoff delay. The handoff only takes place in the overlap region between the two APs. Since the handoff detection mechanism employed in FMIP uses signal strength from the beacons received from APs in the vicinity of the MN, the effect of range between the two APs has minimal effect. This is because the MN node will only initiate handover if, and when, the beacon it receives from another AP other than the current one is stronger. This takes place in the overlap region and thus it is the extent of the overlap that affects the handover rather than the range of APs. From a micromobility perspective, the integrated scheme and FMIPv6 relatively suffer from the same average handover delay as the WLAN AP range changes. The figure shows that range has marginal effect on handover latency in both schemes and therefore, no significant change in handover latency was observed. On a small-scale network, WLAN configurations do not affect the overall latency delay for the protocol schemes. From the simulation results, the performance of realtime applications was not adversely affected in the integrated scheme as compared to pure SIP scheme due to shorter disruption time and lower packet loss. The integrated scheme offered smooth handover resulting in lower packet loss with minimal effect on the VoIP application.

6. CONCLUSION

An integrated fast Mobile IPv6 and SIP handover management mobility architecture is proposed that exploits both the complementary capabilities of each protocol and aims at reducing their functionality redundancies. The basic idea in the mobility framework is to support various mobility scenarios by making use of FMIPv6 and SIP procedures in a jointly optimized way to improve performance.

From the simulation results, the proposed mobility architecture can offer powerful mobility support in terms of seamless handover to mobile devices for IP-based next-generation networks. The basic idea of the mobility framework has been to jointly optimize the capabilities of the network layer protocol FMIPv6 and the application layer protocol SIP. Thus, the architecture offers flexibility to be adapted in future network developments to support realtime applications effectively under the "always best connected concept."

ACKNOWLEDGMENTS

This work is supported in part by Telkom, Nokia Siemens Networks, TeleSciences, and National Research Foundation, South Africa, under the Broadband Center of Excellence program.

REFERENCES

- [1] R. Koodli, Ed., *Fast Handovers for Mobile IPv6*, IETF Internet Draft, draft-ietf-mishop-fast-mipv6-03.txt, October 2004.
- [2] J. Rosenberg, H. Schulzrinne, G. Camarillo, et al., "SIP: Session Initiation Protocol," IETF RFC3261, June 2002.
- [3] E. Wedlund and H. Schulzrinne, "Mobility support using SIP," in Proceedings of the 2nd ACM/IEEE International Conference on Wireless and Mobile Multimedia (WOWMOM '99), pp. 76– 82, Seattle, Wash, USA, August 1999.
- [4] H. Schulzrinne and E. Wedlund, "Application-layer mobility using SIP," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 4, no. 3, pp. 47–57, 2000.
- [5] S. Thomson, T. Narten, and T. Jinmei, "IPv6 Stateless Address Autoconfiguration," IEFT Internet Draft, draft-ietf-ipv6-rfc2462bis-07.txt, December 2004.
- [6] S. D. Park, "Minimized Duplicate Address Detection," IETF Internet Draft, draft-park-dna-mdad-00, September 2003.
- [7] T. Guenkova-Luy, A. J. Kassler, and D. Mandato, "End-to-end quality-of-service coordination for mobile multimedia applications," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 5, pp. 889–903, 2004.
- [8] The UCB/LBNL/VINT Network Simulator-ns (version2), http://www.isi.edu/nsnam/ns.
- [9] R. Hsieh, FHMIPv6 Extension for ns2, http://mobqos.ee.unsw.edu.au/~robert/.
- [10] NIST-SIP, http://www-x.antd.nist.gov/proj/iptel/.
- [11] R. Hsieh and A. Seneviratne, "A comparison of mechanisms for improving mobile IP handoff latency for end-to-end TCP," in *Proceedings of the Annual International Conference on Mobile Computing and Networking (MOBICOM '03)*, pp. 29–41, San Diego, Calif, USA, September 2003.