A Robust Handover Management Mechanism based on Host Identity Protocol (HIP)

Muhana Muslam, student member IEEE , H. Anthony Chan, Fellow IEEE, Neco Ventura, Member IEEE and Petro Ernest Pesha, student member IEEE
Electrical Engineering Department
University of Cape Town, P. O. Box 7701, Cape Town Rondebosch
Tel: +27 02 6502804, Fax: +27 02 6503465
email: {muslam, neco, pernest }@ crg.ee.uct.ac.za, h.a.chan@ieee.org

Abstract-The support of macro-mobility and micro-mobility in an efficient, seamless and secure way is one of the key features of the Next Generation Network (NGN). Host Identity Protocol (HIP) is a mobility management solution developed by the Internet Engineering Task Force (IETF). It efficiently supports macro-mobility but has excessive signaling and long handoff latency in micro-mobility scenario. This paper introduces a new technique to efficiently support micro-mobility based on HIP capabilities while providing the same security level as the HIP. Performance analysis shows that this new method is more efficient than existing ones.

Index Terms—HIP, Mobility, Micro-mobility

I. Introduction

As the Internet becomes increasingly mobile and overloaded with mobile communication devices, mobile users are bound to roam freely and attach to a variety of networks. Host mobility becomes one of the key features of the Next Generation Network (NGN) which is the All-IP based heterogeneous networks [1]. Since the standard Internet design is based on stationary hosts, the host mobility in the Internet introduces some technical challenges, such as session continuity, host reachability, and security threats. Addressing these challenges might seem straightforward, but there are many issues that make them very complex, such as the duality problem of IP addresses in the standard Internet [2]. The duality problem refers to that the IP address serves as both host identifier and locator at the same time in the Internet.

When a Mobile Node (MN) is moving during a communication session it might move within a single domain or move outside the current domain to a totally different domain. The first case is referred to as "micro-mobility", which is the area this paper investigates. The latter is referred to as "macro-mobility" [3]. These two main scenarios can be managed at different layers of the conventional TCT/IP stack [4]. Since the IP is ubiquitous in the Internet, the IP layer solutions are most general. Mobile IP (MIP) [5], which is one of the IP layer solutions, extends the ability of the Internet to support the host mobility but it still has security threats that have prevented the deployment of IP mobility extensions. Due to these security threats such as Denial of Service attack (DoS), services provided by Internet servers will frequently be unavailable [6].

Therefore, it is necessary to deploy additional mechanisms to solve the security threats. The addition of such mechanisms to ensure efficient and secure mobility in the conventional TCP/IP stack is very complex. It is complex because the conventional TCP/IP is already overloaded with previously added functionality [6]. This has motivated researchers to find a new approach to efficiently support mobility and provide a new basis for future mobile Internet. Host Identity Protocol (HIP) is one of the new approaches developed by Internet Engineering Task Force (IETF) in the mobility and security area. It provides mobility support in a simple and elegant way than other proposed solutions [6]. Additionally, references [7, 8] show that HIP is more elegant than MIP which is one of most popular mobility management solution.

HIP serves well as a macro-mobility management solution but introduces long handover latency and unnecessary control messages in a micro-mobility environment [9]. Such long handover latency increases the probability of packet lost and delay as well as decreases the performance of the upper layer application, particularly the real time application such as Voice over IP (VoIP). Therefore, there is need to develop an adequate and efficient solution to reduce handover latency as well as mobility-related signaling of HIP in micro-mobility environment while retaining the same security level of the HIP. The proposed solution should support mobility to allow the mobile users access their services wherever they go in a secure and efficient manner. The rest of the paper is organized as follow. We briefly describe HIP in the next Section. In Section III, we discuss the related work and our micro-mobility solution is presented in Section IV. This is followed by performance analysis in Section V and conclusion in Section VI.

II. Host Identity Protocol (HIP)

The IP duality problem of current IP scheme is one of the main issues that make mobility management very difficult [2]. In the basic IP-based networks, the IP address is used both as an identifier of a network session and as a topological location determiner. The mobile node obtains a new IP address whenever it moves to a different network. So the upper layer protocols such as TCP and UDP which are bound to the IP addresses need to be informed about the new IP address to maintain the ongoing session. This update normally leads to reconfiguration of already established state which may result in disconnection of the ongoing session. Host Identity Protocol (HIP) [10], which is a candidate

protocol for a secure mobile Internet, is an alternative and elegant solution to the IP duality problem. It introduces a new namespace (host identity name space) to serve as host identifier and uses the current IP addresses as locators. Thus the identifier is used to establish and maintain a communication session between the communicating parties while the locator identifies the current point of attachment of the host (e.g. MN). HIP introduces a new sub-layer into IP/TCP stack, which decouples the transport layer from the inter-networking layer making the ongoing communication independent of the host location. Figure 1 depicts the HIP protocol stack.

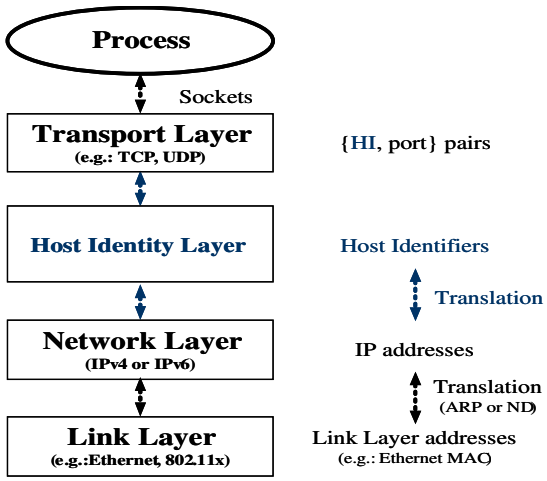

Figure 1. The HIP Protocol stack.

A. *HIP Base Exchange*

HIP operation works in two modes, which are the control mode (Base Exchange) and the data traffic mode. In the HIP Base Exchange (BE), the communicating parties establish a pair of security association between themselves [10]. The BE consists of four messages that are exchanged in two round-trip times. Figure 2 describes the establishment of the security association between two communicating parties, each having a single IP address.

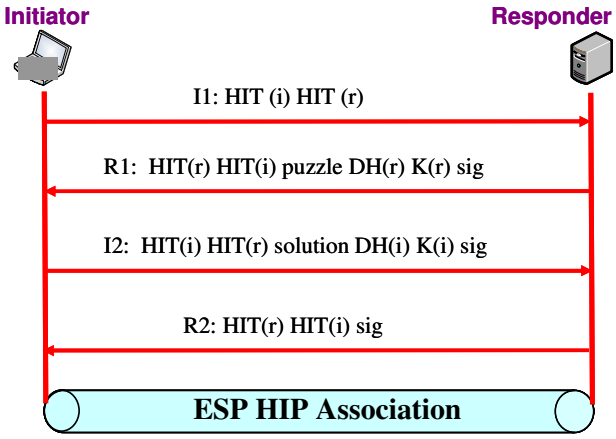

Figure 2. The HIP Base Exchange Protocol.

B. *Mobility and multihoming in HIP*

Mobility and multihoming support for HIP is presented in a separate IETF document [13]. This extension allows the HIP host to notify its peers about the set of new locators by exchange of the three UPDATE messages. When an MN moves, it includes the new locator (i.e., a more general concept than an IP address), into an UPDATE packet and sends to its peers. Thus the peers of the MN can redirect the traffic to the new location of the MN after the required address check is performed. The purpose of the address check is to protect against different security threats due to mobility such as a Denial of Service attack (DoS) and traffic flooding.

III. Related Work

This section briefly describes the existing micro-mobility solutions for HIP and their shortcomings. The handover latency varies in different handover scenarios [14], for example, in macro-mobility scenario and in micro-mobility scenario, the latter being the main concern of our study.

There is no adequate micro-mobility management solution for HIP but some proposed solutions [15, 16, and 17]. Work has been done on micro-mobility architecture for HIP that uses a local rendezvous server (LRVS) [15]. The LRVS extends the concept of the normal HIP rendezvous server (RVS). It performs Network Address Translation (NAT) as well as the normal RVS functions. Once the MN enters a given local domain, it detects the LRVS in the visited network either by actively initiating a service discovery procedure or passively waiting for a service announcement according to the extension specified in "HIP Service Discovery – Internet Draft" [18]. Then the MN registers itself at the LRVS using the registration extension specified in [19]. The LRVS also registers its IP address and the MN information at the Domain Name Server as specified in [20]. The MN, therefore, notifies the LRVS instead of the correspondent node (CN) to redirect the data traffic to its new location, i.e. the new local IP address. However, this solution does not avoid the IP address configuration and re-registration at the LRVS whenever the MN moves from one subnet to another within the same domain. The IP configuration, which in turn required Duplicate Address Detection (DAD), adds some delay to the handover process while the re-registration takes considerable time that also contributes to the handover latency. Moreover, for the registration, the MN always sends its new IP address to the LRVS even when there is a cross-over point between the old point of attachment and the new one. So the required time to do the re-registration at the LRVS is relatively higher than to do it at a topologically closer one (i.e., cross-over). Furthermore, before the registration at LRVS is over, the LRVS forwards the packets that are destined to the MN to the old Access Router (AR) of the MN. This increases the probability of packet delay and loss. Figure 3 illustrates the signaling flow of the scheme.

The second method is an approach for securing micro-mobility that is more reasonable for hierarchical mobility domains [16]. It focuses on the authentication of the location binding update messages to prevent the possible security issues such as man-in-the-middle attack and DoS. It uses regional anchor point (RAP) which supports the dynamic binding between the end-point identifiers (EIDs)

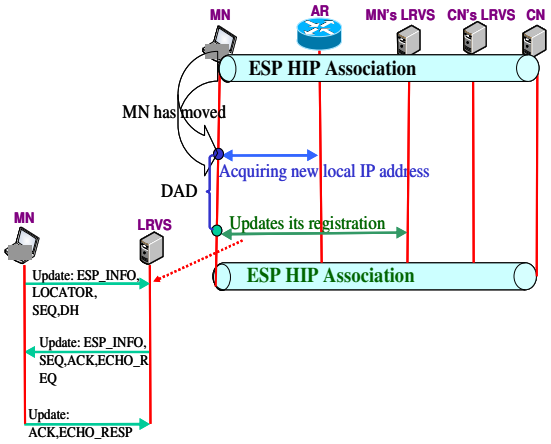

Figure 3. Intra-domain handover of Novaczki et. al.

and their IP addresses. Once an MN enters a given region, it does not need to register at any mobility anchor points within the given region. During the security association establishment, mobility anchor points in the region learn the required security context and current location information of the MN, while the nearest anchor point only knows the shared secret key of the communication session. The solution is based on the use of Lamport one-way hash chains and secret splitting techniques to bind the messages of location updates together and to establish a security association (SA) between the MN and the nodes (e.g. anchor points in a domain) along the path of the MN and CN.

However, this solution is not a good micro-mobility management solution because it behaves as a macro-mobility solution in many situations. For example, if the Lamport one-way hash chain reaches the seed value or a man-in-the-middle attack between the MN and nearest anchor point (N-AP) occurs, this scheme behaves as macro-mobility solution and also requires the creation of a new hash chain. Furthermore, the scheme still needs to reconfigure its local IP (LIP) if the MN changes its point of attachment, thus affecting the handover latency, signaling overhead and packet loss as well as compromising location privacy. The signaling flow is illustrated in Figure 4.

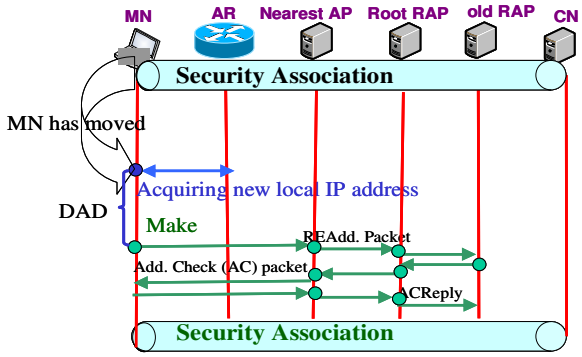

Figure 4. Intra-domain handover signaling of Ylitalos *et. al.*

Finally, a HIP based mobility management architecture scheme which uses tight coupling between the UMTS and WLAN is proposed in [17]. The architecture uses a rendezvous server (RVS) in the UMTS network to handle the handover process with a strategy to establish a new connection before terminating the previous one. However, it still suffers from the same problem that is faced by [15, 16] in terms of IP configuration delay as a result of IP address changes. The signaling flow of this scheme is similar to the signaling flow of the scheme in [15], but they use the RVS to manage the mobility in a domain rather than using the LRVS. Even though the handover performance is improved, it still needs to be optimized.

IV. Proposed Solution

We propose a HIP micro-mobility solution in which we utilize the advantages of keeping the IP addresses of the Mobile Nodes stable in a given domain. The mobility entities in our architecture scheme are responsible for tracking the movements of the host and the exchange of the required mobility signaling on behalf of the mobile node. The core functional entities are the Local Rendezvous Server (LRVS) which is proposed in [16] and the subnet-Rendezvous Server (S-RVS) which we introduce. The LRVS is responsible for maintaining the mobile node's "reachability" information. The S-RVS is the entity that performs the mobility related signaling on behalf of a mobile node. It resides on the access link where the mobile node is attached. It is also responsible for detecting the movements of mobile node to and from the access link and for initiating the update messages to the nearest anchor point (N-AP). N-AP is the cross-over point between the old point of attachment (PoA) and the new one of that mobile node. The architecture of our method is shown in Figure 5.

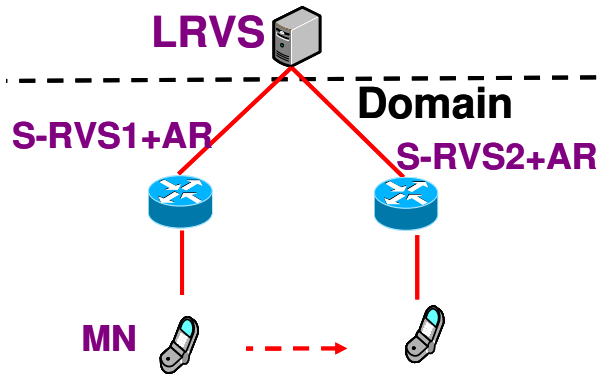

Figure 5. The proposed Architecture.

The design of our scheme is based on the following principles

- Distributed caches are used to store fresh R1 pre-computed packets.
- MN will use the same IP address as it remains within a single domain.

The caching of fresh pre-computed R1 packet at LRVS optimizes the handover performance when re-keying is required due to a handover.

Figure 6 shows different subnets, i.e., different wireless access networks connected to the Internet through a Local Rendezvous Server (LRVS), which manage the mobility within a given domain. We assume that the MN and CN are

registered at the RVS, which is outside the domain managed by the LRVS. The CN is a fixed HIP host in a different domain.

When a Mobile Node (MN) enters a given domain, the complete process is illustrated in Figure 6 INITIAL REGSTRATION part and explained hereafter. The S-RVS on the access link to which the MN is attach to, i.e., S-RVS1 after detecting the mobile node sends an UPDATE packet (UPDATE packet 1) with registration flag to the LRVS. The packet including the host identity tag (HIT) of the MN. Upon reception of UPDATE packet 1, the LRVS sends UPDATE packet 2 to both S-RVS1 and S-RVS2. This may ensure that the network prefix will be delivered to that MN at any subnet in the given domain. The S-RVS1 on receiving of the UPDATE packet 2 sends a Router Advertisement (RA) including the network prefix. Then the MN configures its IP add using the network prefix.

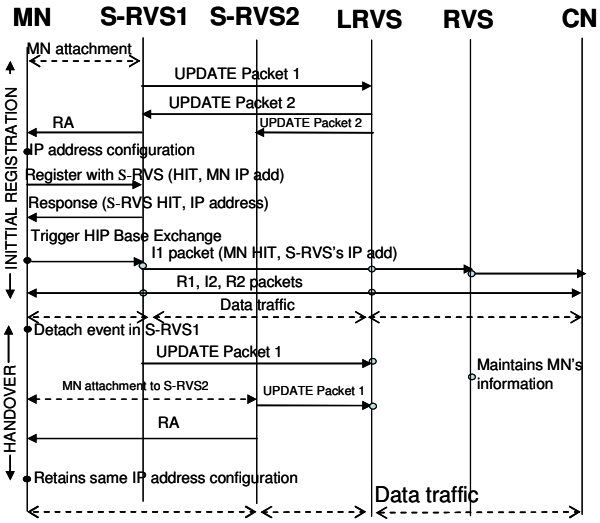

Figure 6. Registration and Intra-domain handover mechanism.

If the MN wants to communicate with a HIP host (e.g. CN), it needs to first establish a HIP security association (SA) as is explained in II A. In this case, the MN sends I1 packet to the S-RVS1, which in turns sends it to the LRVS. The LRVS forwards the packet to the CN through the RVS. The exchange of remaining packets (i.e., R1, I2, and R2) goes directly between the MN and CN. After successfully exchanging these packets, a HIP SA will be established between the initiator (i.e., MN) and responder (i.e., CN). Through the SA establishment all the S-RVSs along the path between the MN and the LRVS will be aware of the SA context.

When the MN performs intra-domain handover, S-RVS1 (i.e. old S-RVS) detects the detachment of the MN and sends UPDATE packet 1 to the LRVS to de-register its IP address i.e., to tell the LRVS that it is no longer the serving S-RVS. The new S-RVS detects the attachment of the MN and sends at the same time an RA and UPDATE packet 1 to the MN and LRVS respectively. The RA includes the same network prefixes that is given to that MN during the registration. So the MN retains the same IP address configuration. This may significantly reduce the handover latency and signaling overhead.

The proposed solution provides a novel structure for HIP micro-mobility management, in which the handover latency and the signaling load are reduced by allowing an MN to use the same IP address (to avoid the duplicate address detection (DAD) process as it remains within a single domain) and sending the MN's HIT and assigned network prefix to the other S-RVSs (e.g. S-RVS2) before a handover occurs. Then the new S-RVS sends at the same time both the same network prefix to the MN and UPDATE packet 1 to the LRVS to set up a new path for the ongoing traffic. The use of the same IP address enables the HIP host to use the established HIP associations during and after the handover since the communication context remains the same and supports the location privacy. Furthermore, reducing the time taken to perform Base Exchange (BE) can further reduce the handover delay when the re-establishment of HIP SA is required. The existence of the S-RVS also eliminates the need for a new location reachability check between the MN and its peer because the new location of the MN is known to the serving S-RVS. The number of S-RVSs in a domain depends on the domain's size and each subnet is managed by an S-RVS which acts as the authoritative S-RVS for that subnet. When the authoritative S-RVS is not functional, its subnet will be managed by a neighboring S-RVS. The number of MNs in each subnet must not exceed the capability of the S-RVS. Using this method, the proposed scheme can manage the "double jump" scenario (i.e., when the communication parties move at the same time) and also multihoming in an easy and efficient way.

V. PERFORMANCE ANALYSIS

In this section we briefly compare the handover latency involved in [14, 15, and 16] and our proposed scheme. The analysis focuses on the delay in the network transmission and latency due to the IP configuration whenever the MN moves. The latency related to the processing of the location binding update messages is not considered in this model. We use the Round Trip Time (RTT) as the measure for the propagation delay. One RTT is defined as the time required for the source to transfer one packet to and from the destination. $RTT_{A,B}$ represents the time required for the packet, sent by A, to reach B as well as back from B. In addition, the RTT is based on the distance the packets traveled from the source to the destination, for example the number of hops between the source and the destination. Our method and all related work [15, 16, and 17] are aimed at providing a better handover performance than HIP for the micro-mobility environment. All schemes aim to reduce the signaling between the mobility anchor point of the MN and its peers. For example, the CN and RVS are not notified of the intra-domain handover action.

The following section explains the handover latency involved in [15, 16, and 17] and our proposed scheme.

A. *Scheme of Novaczki et. al.*

When a MN performs a handover from a previous AR to a new one, the following handover latency components are involved;

• The latency due to the IP address configuration of the MN at the new point of attachment, LIP-CONFG.

• Latency due to the three way handshake readdressing protocol, 1.5RTTMN, LRVS, which depend also on the distance between the MN and the LRVS.

Thus, the handover latency for a Novakzki's scheme mobility management protocol is shown below.

$$L_{Novaczki} = L_{IP\text{-}CONFIG} + 1.5\ RRT_{MN,\ LRVS} \quad (1)$$

Note that (1) also reflects the handover performance of reference [17], but the difference is the exchange of location update messages done between the MN and the RVS which is located with the gateway in [17].

B. *Scheme of Ylitalo et. al*

When a mobile node moves from one PoA to another, an MN exchanges the readdressing messages with a nearest anchor point (N-AP). The following handover latency components are involved;

• The latency due to the IP configuration of the MN at the new point of attachment, LIP-CONFIG.

• Latency due to the three way handshake readdressing protocol, 1.5RTTMN, N-AP, which depend also on the distance between the MN and the N-AP.

Thus, the handover latency due to mobility management protocol in this scheme is shown below.

$$L_{Yiltalos} = L_{IP\text{-}CONF} + 1.5\ RTT_{MN,\ N\text{-}AP} \quad (2)$$

Note that the handover latency in (2) is lower than that in (1) because the distance between the MN and the N-AP is shorter than the distance between the MN and the LRVS in the best case (i.e. the N-AP $\neq$ LRVS).

C. *Our scheme*

The proposed scheme overcomes the shortcomings of Yilitalo's and Novakzki's solutions to efficiently manage mobility in a localized domain. Our scheme has the following components contributing to handover delay.

• Latency due to the two way handshake readdressing protocol, RTTS-RVS, N-AP, which depend also on the distance between the S-RVS and the N-AP.

Thus, the handover latency due to our scheme is shown below.

$$L_{Ourscheme} = 0.5\ RTT_{S\text{-}RVS,\ N\text{-}AP} \quad (3)$$

It should be noted that in our scheme, the MN uses the same IP address, so there is no need to update the established HIP association during and after handover since the communication context remains the same. The security and authentication aspects are maintained by the hash chain proposed in [16]. Let us denote the distance between the MN and LRVS in [15] by X, the distance between the MN and N-AP in [16] by Y, and the distance between the S-RVS and N-AP in our scheme by Z. Naturally Y and Z will be equal and both of them are shorter than X. Additionally, (3) shows lower handover latency than (1) and (2), see Figure 7. The figure shows the comparison between our scheme and other schemes that are mentioned in the related work section. The senders of the Location Update (LU) messages in every scheme are placed at the LU sender column in the figure. The responders of these messages are placed in positions that relatively represent their distance from the senders.

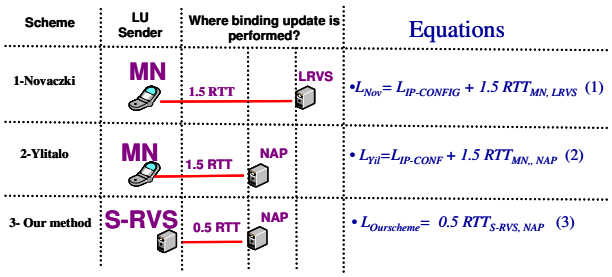

Figure 7. Handover Performance Summary of the different Schemes.

VI. Conclusion

We have proposed a new method that is aimed at optimizing the handover performance for HIP in a micro-mobility environment. We have introduced a new mobility functional entity, which is called the Subnet Rendezvous Server (S-RVS). The S-RVS tracks the mobile node movement and acts on its behalf whenever it moves. This reduces the handover latency and signaling overhead which in turn reduces packet loss and delay. The proposed scheme will also supports location privacy as well as optimizing the handover latency when re-keying is required. The performance analysis shows that proposed scheme is expected to outperform the other proposed micro-mobility solutions presented in [15, 16, and 17]. The modeling and implementation of the proposed scheme is ongoing using the OMNet++v4.0 simulator.

References

[1] F. M. Chiussi, D. A. Khotimsky, and S. Krishnan, "MobilityManagement in Third-Generation All-IP Networks,"IEEE Commun. Mag., Sept. 2002, pp. 124–35.

[2] P. Pacyna, "Advances in Mobility Management for the NG Internet," In China Communications, vol. 3, Issue 3, June 2006.

[3] Pierre Reinbold, (FUNDP) and Olivier Bonaventure, (UCL), "IP Micro-Mobility Protocols," IEEE Communications Surveys & Tutorials Third Quarter 2003, pp 40-57.

[4] Deguang Le, Xiaoming Fu, Dieter Hogrefe, "A review of mobility support paradigms for the internet," IEEE Communications Surveys & Tutorials, V. 8, I. 1, pp. 38-51, 2006.

[5] C. Perkin, "Mobile Networking through Mobile IP," IEEE Internet Computing, Volume 2, Issue 1:58-69, Jan.-Feb. 1998.

[6] A. Gurtov, Host Identity Protocol (HIP): Towards the Secure Mobile Internet, ISBN 978-0-470-99790-1, Wiley and Sons, June 2008. (Hardcover, 332 p).

[7] A. Khurri, E. Vorobyeva, A. Gurtov, "Performance of Host Identity Protocol on Lightweight Hardware," in Proceedings of ACM MobiArch, August 2007.

[8] P. Jokela, T. Rinta-aho, T. Jokikyyny, et al., "Handover performance with HIP and MIPv6," 1st International Symposium on Wireless Communication Systems, 2004, vol. 3, pp. 324 - 28, 2004.

[9] T. R. Henderson, J. M. Ahrenholz, and J. H. Kim, "Experience with the Host Identity Protocol for Secure Host Mobility and Multihoming," IEEE Wireless Communications and Networking, V. 3, pp. 2120-2125 March 2003.

[10] R. Moskowitz and P. Nikander, "Host Identity Protocol (HIP) Architecture," IETF, RFC4423, May 2006.

[11] P. Jokela, R. Moskowitz, and P. Nikander, "Using the Encapsulating Security Payload (ESP) Transport Format with the Host Identity Protocol (HIP)," IETF, RFC 5202, April 2008.

[12] J. Laganier and L. Eggert, "Host Identity Protocol (HIP) Rendezvous Extension," IETF, RFC 5204, April 2008.

[13] P. Nikander and T. Henderson, "End-Host Mobility and Multihoming with the Host Identity Protocol," IETF, RFC 5206, April 2008.

[14] S. Yankov and S. Wiethoelter, “Handover Blackout Duration of Layer Mobility Management Schemes,” TKN Technical Report,Berlin, May 2006.

[15] S. Novaczki, L. Bokor, and S. Imr e, "Micromobility support in HIP: survey and extension of host identity protocol," 2006, pp. 651-54.

[16] Ylitalo, J. Melén, P. Nikander, et al., "Re-thinking Security in IP-based Micro-Mobility," in Pro. Of the Informational Security Conference (ICS'04), Palo Alto, CA, USA, 2004, pp. 318-29.

[17] J. So and J.Wang., “A HIP-based Mobility Management for UMTS/WLAN Integrated Networks,” Australian Telecommunication Networks and Applications Conference (ATNAC 2006), 2006.

[18] P. Jokela, J. Merlen, J. Yilitalo,"HIP Service Discovery," IETF Internet Draft <draft-jokela-hip-service-discoery-00>,june 2006.

[19] J. Laganier, T. Koponen, L. Eggert: "Host Identity Protocol (HIP) Registration Extension," IETF, RFC 5203, April 2008.

[20] P. Nikander and J. Laganier, "Host Identity Protocol (HIP) Domain Name System (DNS) Extension," IETF, RFC 5205, April 2008.

Muhana Muslam received his undergraduate degree in 2003 from Computer Man College (CMC) for computer studies and his Master degree in 2006 from the University of Khartoum, Sudan. He is presently studying towards his PhD in the Electrical Engineering Department at the University of Cape Town (UCT). His research interests include Mobility Management and Wireless Networks.