

Secure Network Mobility (SeNEMO) for Real-Time Applications

Tuan-Che Chen, *Student Member, IEEE*, Jyh-Cheng Chen, *Senior Member, IEEE*, and
Zong-Hua Liu, *Student Member, IEEE*

Abstract—The IETF Network MObility (NEMO) working group has considered how to enable an entire network to move from one location to another. Mobile Virtual Private Network (VPN) has been developed to secure mobile user's communication between untrusted external networks and the protected private internal network. However, the IETF's mobile VPN does not address how to support NEMO. In addition, it is not suitable for real-time applications. In this paper, we propose architecture and protocols to support VPN in NEMO, which is called *Secure NEMO (SeNEMO)*. The proposed SeNEMO, based on Session Initiation Protocol (SIP), is specifically designed for real-time applications over VPN. It allows an entire network to move and still maintains session continuity. In addition to analyzing the security vulnerabilities, we also propose analytical models to evaluate the performance of the proposed SeNEMO. The analysis is validated by extensive simulations. The results show that the proposed SeNEMO can reduce signaling cost significantly.

Index Terms—Network mobility (NEMO), mobile virtual private network (VPN), security, session initiation protocol (SIP), performance analysis.

1 INTRODUCTION

As more and more electronic devices are equipped with wireless communication interfaces, wireless Internet access is not limited to a single device but also a group of devices moving altogether. For instance, a person carrying a laptop and other devices with different radio interfaces such as 3G/4G, WiFi, or Bluetooth may constitute a simple *mobile network*. The laptop can be a mobile router connecting to the Internet. All other devices can obtain Internet access through the laptop. Therefore, *Personal Area Networks (PANs)* can be an example of mobile network. Other examples of mobile networks include wireless networked devices and sensors deployed inside vehicles such as buses, cars, and trains (e.g., FIFTH project [1]). In addition, spacecrafts and airlines (e.g., WirelessCabin project [2]) are deploying mobile networks for Internet access and control systems.

The IETF Network MObility (NEMO) working group has completed several RFCs to enable a network to move from one location to another location while still maintaining its local nodes' ongoing sessions. In the NEMO basic support protocol [3], a mobile network may contain several nodes which are connected by one or more Mobile Routers (MRs). An MR, which connects to the IP backbone, is in charge of the mobility management of the entire network. Mobile Network Nodes (MNNs) in NEMO are classified as: Local Fixed Nodes (LFNs) and Visiting Mobile Node

(VMNs). An LFN always connects to the same mobile network, while a VMN can change its point of attachment. Similar to that in Mobile IP (MIP) [4], every *mobile network* is assigned Mobile Network Prefixes (MNPs) by the home network. The Home Agent (HA) needs to maintain a binding between the MNPs and the MR's current Care of Address (CoA). Therefore, when the mobile network moves away from its home network, packets from Correspondent Nodes (CNs) are redirected by HA to the MR. The acronyms used in this paper is listed in Table 1.

In this paper, we consider how to provide Virtual Private Network (VPN) services in NEMO. Security has become a critical issue for today's Internet. VPN has been developed to secure user's communication between untrusted external networks (internet) and the protected private internal network (intranet). VPN services over NEMO can be used in a variety of applications so a mobile network can access to its intranet in a secure way. For example, a NEMO VPN can be used in public safety, where wireless devices in a police patrol car can access to the criminal databases, driver license and vehicle registration databases, or other services in the dispatch center as the car travels between different subnets. Similar type of services can also be used in ambulance or mobile medical car, where various wireless devices or sensors are deployed inside the car. In addition, NEMO VPNs can also be adopted by enterprise mobile users who carry several wireless devices and are required to maintain secure sessions to their company's intranet. However, the IETF NEMO working group has been concluded in June 2008. How to provide VPN services were not solved in IETF NEMO working group. Although IETF has proposed a VPN architecture to support mobility [5], the solution is for a single node only. In addition, it is based on MIP, which is not suitable for real-time applications. As discussed in [6], [7], the IETF's mobile VPN employs one IPsec [8] tunnel and two MIP tunnels. The three tunnels cause large overhead for transmitting real-time packets. In this paper, we propose

• T.-C. Chen and Z.-H. Liu are with the Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan 300.
E-mail: {brad.tcchen, horselui}@gmail.com.

• J.-C. Chen is with the Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan 300. E-mail: jchen@ieee.org.

Manuscript received 9 July 2009; revised 23 Jan. 2010; accepted 12 Aug. 2010; published online 15 Nov. 2010.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-2009-07-0279. Digital Object Identifier no. 10.1109/TMC.2010.219.

TABLE 1
List of Acronyms

ALG	Application Level Gateway	RTCP	RTP Control Protocol
AVP	Attribute Value Pair	RTP	Real-time Transport Protocol
CN	Correspondent Node	SA	Security Association
CoA	Care of Address	SDP	Session Description Protocol
HA	Home Agent	SeNEMO	Secure NEMO
IKE	Internet Key Exchange	SIP	Session Initiation Protocol
i-HA	internal HA	SIP-NVG	SIP NEMO VPN Gateway
i-MIP	internal MIP	SRTP	Secure Real-time Transport Protocol
MAA	Multimedia-Auth-Answer	TEK	Traffic Encryption Key
MAR	Multimedia-Auth-Request	TGK	TEK Generation Key
MIDCOM	Middlebox Communication	UAA	User-Authorization-Answer
MIKEY	Multimedia Internet Keying	UAR	User-Authorization-Request
MIP	Mobile IP	VPN	Virtual Private Network
MN	Mobile Node	VPN-TIA	VPN Tunnel Inner Address
MR	Mobile Router	x-HA	external HA
NEMO	Network Mobility	x-MIP	external MIP
PAN	Personal Area Network		

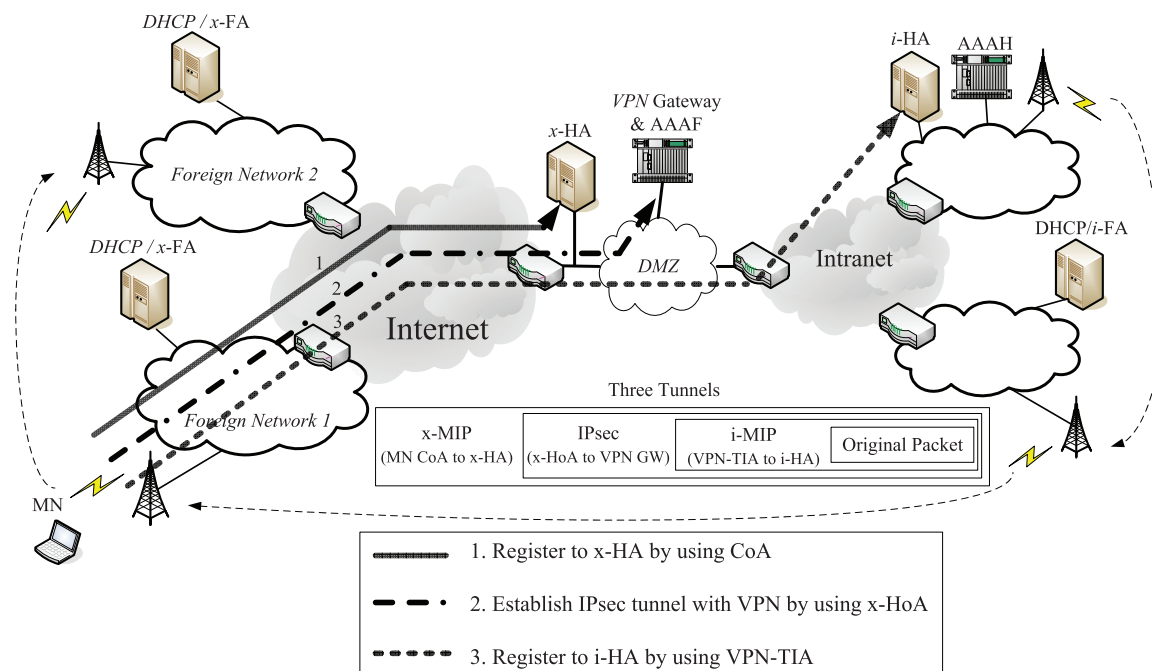


Fig. 1. MVPN proposed by IETF.

architecture and protocols to support VPN in NEMO, which is called *Secure NEMO (SeNEMO)*. The proposed SeNEMO, based on Session Initiation Protocol (SIP) [9], is specifically designed for real-time applications over mobile VPN for group mobility. The contribution of this paper is two-fold. First, we propose a new architecture and protocols to support VPN in NEMO. We demonstrate that by integrating SIP-based mobile VPN with NEMO, we can provide secure and efficient group mobility for real-time services. Second, we develop an intricate analytic model to evaluate the performance of the proposed SeNEMO. The mathematical model quantifies the performance of the proposed SeNEMO and IETF's mobile VPN. Extensive simulations have also been conducted to validate the analysis.

The rest of this paper is organized as follows: Section 2 provides an overview of related work. The proposed SeNEMO is presented in Section 3. The analytical model and numerical results are presented in Section 4 and Section 5, respectively. Section 6 summarizes this paper.

2 RELATED WORK

As aforementioned discussion, IETF has defined architecture and protocols for mobile VPN [5]. Fig. 1 illustrates the IETF Mobile VPN (MVPN). There are two HAs, internal HA (i-HA) and external HA (x-HA), located in intranet and internet, respectively. When the Mobile Node (MN) moves out of the intranet, it first acquires a new CoA from DHCP

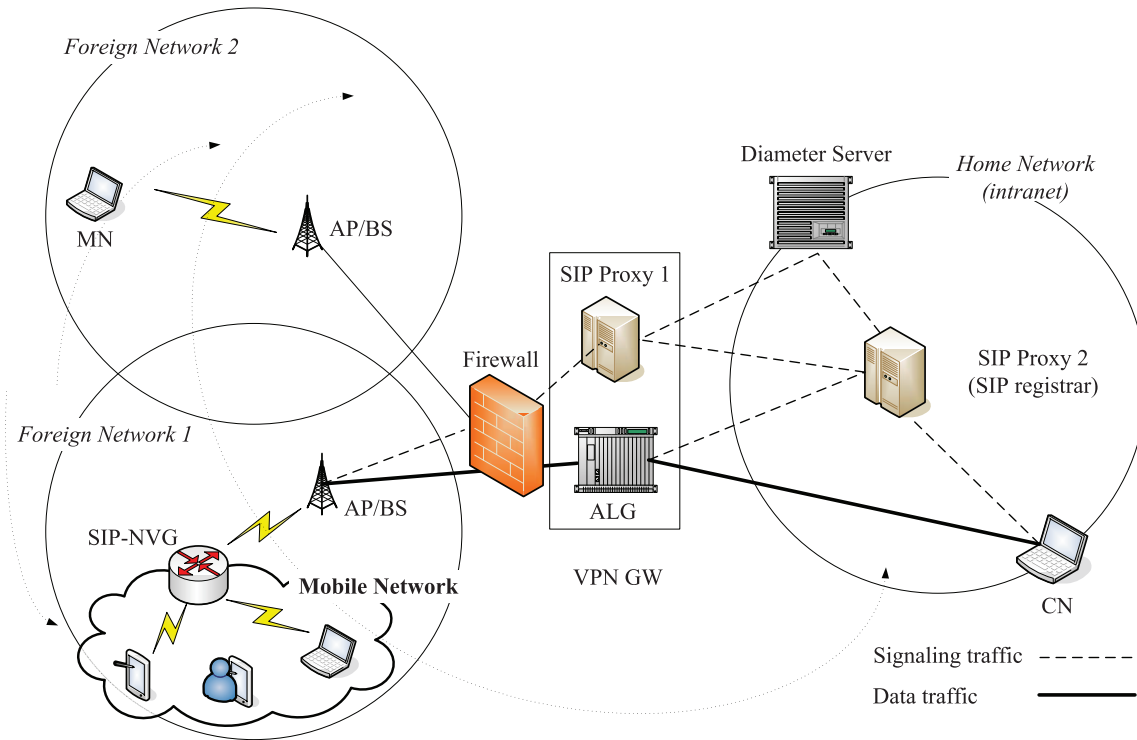


Fig. 2. System architecture.

server or Foreign Agent (FA) and registers the CoA with the x-HA. The MN then establishes an IPsec tunnel with the VPN gateway by using its external Home Address (x-HoA). The IPsec tunnel is built by using Internet Key Exchange (IKE) [10]. A VPN Tunnel Inner Address (VPN-TIA) will be assigned to the MN as the MN's internal CoA during the IKE negotiation. After that, the internal MIP (i-MIP) registration is performed by registering VPN-TIA with its i-HA. With the external MIP tunnel, the IPsec tunnel will not break when the MN roams in internet because only the external MIP (x-MIP) registration is required. The three tunnels (x-MIP, IPsec, i-MIP) is shown in the bottom of Fig. 1.

The IETF MVPN cannot be applied to NEMO because it does not consider the mobility of a group of mobile devices. Besides, the IETF solution, which is based on IPsec and MIPv4 [4], will incur long handoff latency and end-to-end latency [11], [12]. The three tunnels (one IPsec tunnel and two MIP tunnels) increase around 100 bytes for the packet length, which is relatively five times of the length of a G.729¹ real-time packet [13]. The tunnels increase massive overhead in terms of packets length and processing time. This may degrade the performance of real-time applications, which are sensitive to bandwidth and delay. In addition, where to put the x-HA and how to trust the x-HA are critical issues. To better support real-time applications, a SIP-based mobile VPN has also been proposed [6], [7]. However, it considers the VPN for the movement of a single node only.

Although security in NEMO has been investigated in some IETF documents and other literatures [14], [15], all of them are based on MIP and IPsec which inherit the

problems discussed earlier for real-time services. Besides, they are not designed specifically to support mobile VPN. On the other hand, SIP has been proposed to provide host mobility and session continuity [16], [17]. It has also been shown that SIP-based mobility can be easily deployed and reduce the data transmission delay in NEMO [18], [19]. However, by adopting SIP into NEMO, it may increase signaling cost during handoff due to sending many re-INVITE messages for ongoing sessions. In addition, security and VPN are not addressed specifically in the studies.

In this paper, we integrate SIP-based mobile VPN with NEMO and evaluate its performance. To the best of our knowledge, among the literatures available in public domain, this paper is the first considering how to support mobile VPN in NEMO. Our design is particularly suitable for real-time applications.

3 PROPOSED SECURE NEMO (SENEMO)

The proposed SeNEMO comprises SIP, Secure Real-time Transport Protocol (SRTP) [20], Multimedia Internet KEYing (MIKEY) [21], and Diameter [22] to provide VPN services in NEMO.

Fig. 2 depicts the architecture of the proposed SeNEMO. It shows a mobile network in a Foreign Network (internet) connecting to the CN in the Home Network (intranet). The *SIP NEMO VPN Gateway (SIP-NVG)* shown in the Mobile Network residing in Foreign Network 1 is the gateway of the mobile network to other networks. It follows the SIP standards and manages the traffic between the mobile network and outside world. The *VPN Gateway (VPN GW)* consists of SIP Proxy 1 and Application Level Gateway (ALG). There is a firewall between internet and intranet to prevent external SIP users from getting direct access to the intranet. The SIP Proxy 1 is a SIP proxy server, which

1. There are many different voice codecs used in IP networks, including G.711, G.723.1, G.726, G.728, and G.729. The payload lengths of these codecs are in the range of 20 bytes to 160 bytes. G.729 with payload length of 20 bytes is the most commonly used codec for VoIP applications due to its low bandwidth requirement.

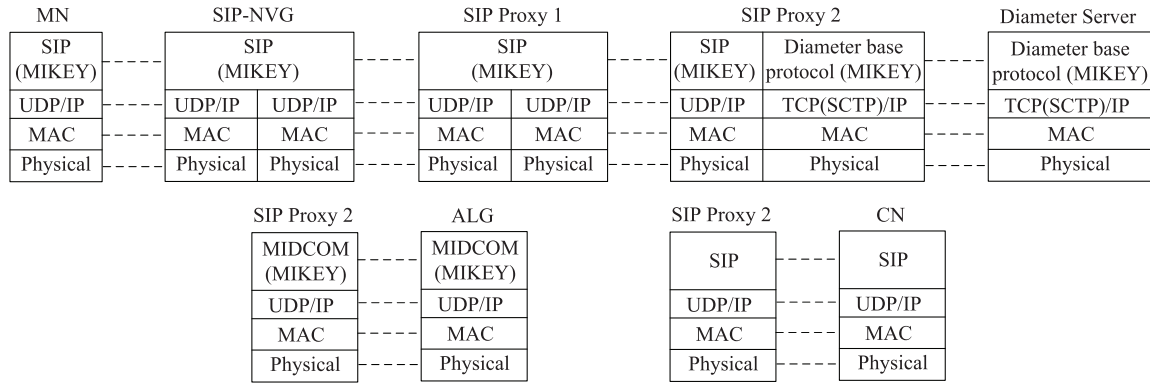


Fig. 3. SeNEMO protocol stacks for control plane when the mobile network is in internet and CN is within intranet.

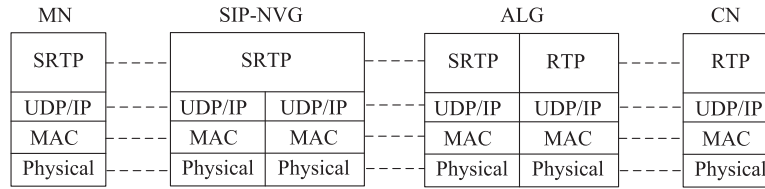


Fig. 4. SeNEMO protocol stacks for user plane when the mobile network is in internet and CN is within intranet.

authenticates the incoming SIP messages through the Diameter server. It also routes messages to SIP Proxy 2 which is essentially a SIP Registrar. Meanwhile, MIKEY is used as key management protocol to provide security keys for the ALG, which then oversees all data traffic.

The main advantages of the proposed SeNEMO are:

- All SIP clients are able to directly communicate with each other without going through a mobility agent such as the HA in MIP. Therefore, the routes are optimized. This is particularly beneficial for real-time applications such as Voice over IP (VoIP) and video streaming. In addition, tunnels such as IPsec tunnel or MIP tunnel are not required.
- When a mobile network changes its point of attachment, a registration request can represent the entire mobile network. This can reduce the signaling overhead significantly.
- The proposed SeNEMO reuses existing protocols. It is compatible with existing standards and is easy to implement. The proposed SeNEMO has been implemented in a testbed [23].

Section 3.1 first describes the protocols used in our proposed architecture. The proposed SeNEMO architecture is then presented in Section 3.2. Section 3.3 discusses the operations of the proposed SeNEMO. Specifically, how to support handoff and maintain session continuity are discussed. In Section 3.4, we analyze the security vulnerabilities of the proposed SeNEMO.

3.1 Protocols

In this section, we discuss the corresponding protocols for *signaling*, *secure transport*, *key management*, and *Authentication, Authorization, and Accounting (AAA)*. Figs. 3 and 4 illustrate the protocol stacks for control plane and user plane in SeNEMO, respectively. For control plane, SIP is the main protocol to manage the session between MN, SIP-NVG, SIP Proxy 1, SIP Proxy 2, and CN. Diameter SIP Application [24], which is based on the Diameter base protocol [22], is

used to authenticate and authorize a user in Diameter Server while the resource allocation in ALG is achieved by using Middlebox Communication (MIDCOM) [25]. In addition, MIKEY messages are embedded inside the messages of Diameter base protocol and Session Description Protocol (SDP) [26] to carry the security information. For user plane, when the mobile network resides in internet, SRTP is used to secure the data transmission between MN and ALG.

3.1.1 Signaling

SIP is an application-layer signaling protocol. It is used to create, modify, and terminate sessions in the proposed SeNEMO. SIP has defined its own security and authentication schemes. In our proposed SeNEMO, we use them to authenticate and identify the mobile users.

SIP also supports *user mobility* and *terminal mobility* [16], [27]. The terminal mobility is achieved by sending new INVITE (re-INVITE) to the CN by using the same call ID as that in the original session. The new INVITE contains the new contact address the MN has acquired in the new location. After receiving the re-INVITE, the CN will redirect future traffic to the MN's new location.

3.1.2 Secure Transport

SRTP defines a framework to provide *encryption* and *integrity* for Real-time Transport Protocol (RTP) [28] and RTP Control Protocol (RTCP) [28] streams. SRTP also provides replay protection based on the RTP sequence number and the index number of RTCP. The predefined cryptographic transformations provide low computational cost and limited packet expansion so bandwidth can be used more economically than IPsec [29]. It is also independent of the underlying transport networks.

3.1.3 Key Management

MIKEY is a key management protocol developed for multimedia real-time applications running over RTP/SRTP. Comparing with IKE which is widely used as key management protocol for unicast, MIKEY is designed for

peer-to-peer or small interactive groups. MIKEY can fulfill the requirements of different environments. For example, MIKEY message can be embedded inside SDP message. A new type k has been defined in SDP to carry MIKEY message.

The main purpose of MIKEY is to transport the *TEK² Generation Key (TGK)* and other related security parameters or policies which are used in security transport protocols. TGK is an upper-level key, which is shared within an interactive group. It is used to derive TEK for each cryptographic session. TEK along with the exchanged security parameters are denoted as the Data Security Association (SA). The Data SA is used as input of security transport protocol, such as SRTP.

3.1.4 AAA

Based on the Diameter base protocol, Diameter SIP Application allows a client of a SIP server to be authenticated and authorized by a Diameter server. There are six Diameter commands in the Diameter SIP application. In the proposed SeNEMO, we use User-Authorization-Request (UAR)/User-Authorization-Answer (UAA) and Multimedia-Auth-Request (MAR)/Multimedia-Auth-Answer (MAA) to process SIP REGISTER and INVITE messages. The authentication is done by Diameter server rather than by delegating to a SIP server.

3.2 Architecture

As aforementioned discussion, the SIP-NVG is the gateway of the mobile network to other networks. When a mobile network roams between different IP subnets, the SIP-NVG not only keeps ongoing sessions unbroken, but also transmits data in a secure manner. Besides, the SIP-NVG ensures that all the MNs attaching to the mobile network are globally reachable at any time.

There are two types of interfaces owned by SIP-NVG: egress interface and ingress interface. A SIP-NVG attaches to internet through *egress interface*. Once a mobile network moves to a new IP subnet, the egress interface of the SIP-NVG will get a new IP address. On the other hand, when an MN wants to join a mobile network, it attaches to the *ingress interface* of the SIP-NVG. In our design, each mobile network has only one SIP-NVG which essentially is an MR with SIP capability. The proposed SIP-NVG is able to route SIP messages and data traffic between its egress interface and ingress interface by translating the corresponding headers.

Fig. 5 depicts the flows for registration when the mobile network is in a foreign network. When an MN enters a mobile network, the MN will get a new IP address and register the new IP address with the SIP-NVG. As shown in Fig. 5, the MN will update its current location with the SIP registrar residing in the home network by sending the REGISTER with the newly obtained contact address. In this example, we assume the mobile network resides in a foreign network and the new address assign for the MN is `mn-1@nemo.vpn.com`. In our proposed architecture, the SIP Proxy 2 not only handles the signaling messages but also acts as the SIP registrar. The registration is based on that specified in Diameter SIP application. As illustrated in Fig. 5, the SIP-NVG translates the *Contact* field in the REGISTER from the MN's address into the SIP-NVG's URI

address, which is `sip-nvg@hs.vpn.com`. Besides, the SIP-NVG establishes a *Mapping Table* to record the registration information for the MN. Hence, each request targeted to the MN will be redirected to the SIP-NVG. The SIP-NVG then can forward the request to the MN according to the Mapping Table.

The proposed architecture depicted in Fig. 2 adopts an ALG which follows MIDCOM architecture. In our proposed architecture, the ALG only accepts commands from the SIP Proxy 2 and provides responses for the corresponding commands. As the data protocol stacks shown in Fig. 4, when the ALG receives a special incoming RTP stream from the home network to an MN in internet, it replaces the whole IP/UDP/RTP headers with a new one, transforms the new RTP packet into SRTP format, and delivers the SRTP stream to the destination. In reverse direction, the ALG receives the SRTP stream from internet. The ALG decrypts it and verifies it to decide whether the SRTP packet is valid or not. If the SRTP packet is decrypted and verified successfully, the RTP payload will be carried by a new RTP header. The new RTP packet is then transmitted to the home network. The symmetric RTP streams are represented as a session in the ALG.

Each session in the ALG requires enough external and internal resources. For example, the external resource includes external listening address, external listening port, external destination address, and external destination port. Destination addresses and ports are provided by the SIP Proxy 2. On the other hand, the internal resource includes internal listening address, internal listening port, internal destination address, and internal destination port. Only when all resources are ready, the session in the ALG will start. When either the external or internal resource is reserved successfully, the ALG will reply the reserved listening address and port to the SIP Proxy 2. The SIP Proxy 2 will replace the original parameters in SDP. Data SA is also a resource which is the input of SRTP. The Diameter server provides the Data SA. The SIP Proxy 2 transmits it with TGK to the ALG in the resource reservation command.

3.3 Operations

In the architecture shown in Fig. 2, the entire mobile network may move altogether from an IP subnet to another. This is referred to as *network handoff*. It is also possible that an MN moves into or moves out of the mobile network. This is referred to as *node handoff*. This section discusses how the proposed SeNEMO maintains a secure sessions during network handoff and node handoff.

3.3.1 Node Handoff

Fig. 6 depicts the flows when an MN moves into a mobile network which is located inside home network. First, the MN registers with the SIP-NVG and the SIP registrar as that discussed in Section 3.2. After the registration, the MN needs to re-invite the CN, if there are active sessions between them. For the INVITE request, in addition to translating the CONTACT field from the MN's address into the SIP-NVG's URI address, the SIP-NVG also adds the *RECORD-ROUTE* field where the SIP-NVG's URI address is inserted. Therefore, the subsequent messages of the existing sessions will be routed by the SIP-NVG. Besides, the SIP-NVG maintains a *Session Table* to record the information of each session, i.e., whether the session is

2. TEK stands for Traffic Encryption Key.

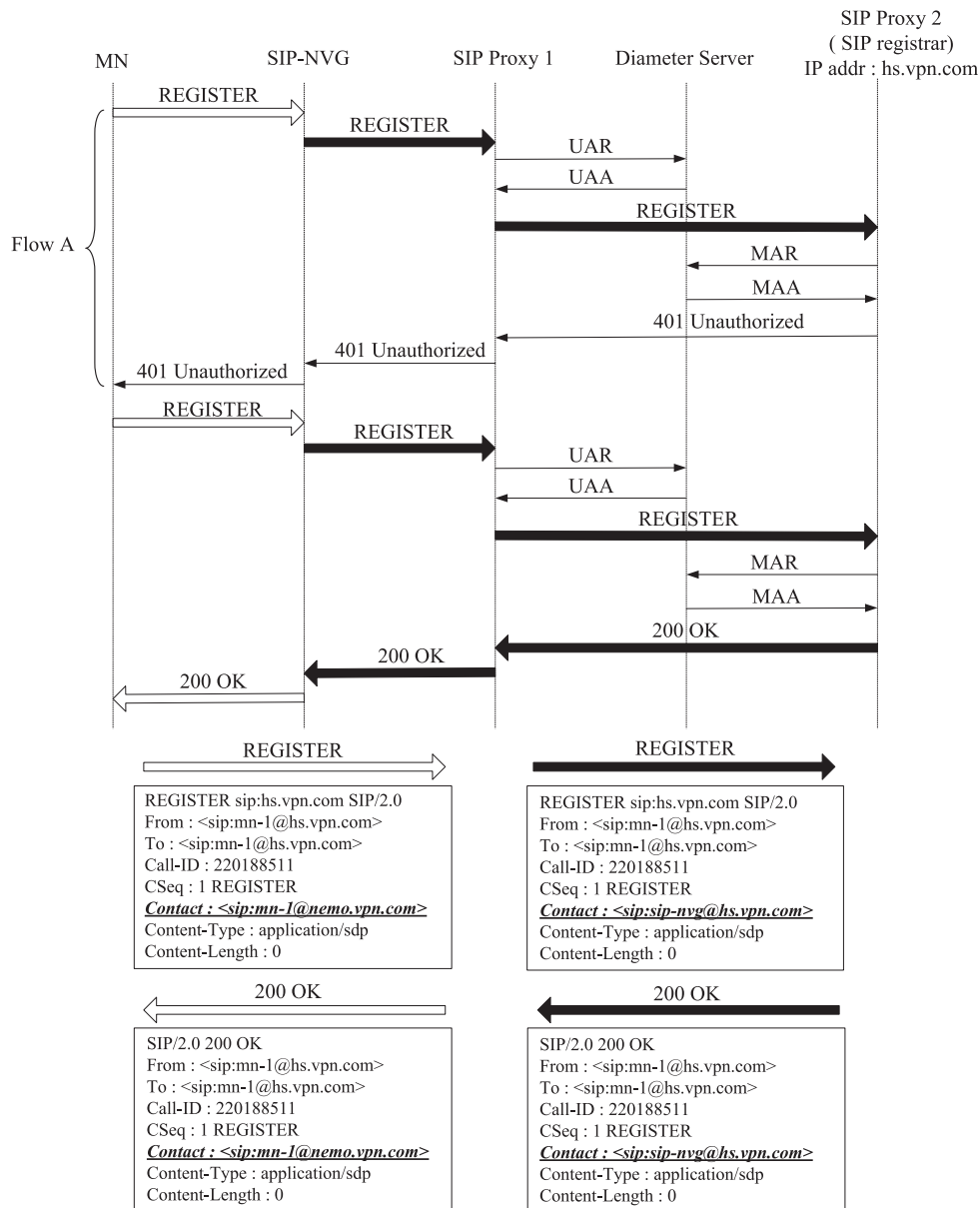


Fig. 5. Message flows and translation of REGISTER when mobile network resides in foreign network.

active and the contact address of the CN. Therefore, the SIP-NVG can determine the ongoing session during handoff.

Fig. 7 depicts the flows when an MN moves into a mobile network which is located inside a foreign network. In this case, the signaling messages will need to go through SIP Proxy Server 1 and be authenticated by the Diameter Server before they reach SIP Proxy Server 2. The first few flows are similar to those in Fig. 6. Rest of them are similar to those in Fig. 8, which will be discussed in next session.

3.3.2 Network Handoff

When the mobile network and the CN are both located inside the same realm, for example, the same intranet or the same network domain in internet, the data traffic between the MNs attaching to the mobile network and the CN does not need to pass through the ALG. Therefore, the SIP Proxy servers in Fig. 2 does not need to intercept SIP signaling messages between the mobile network and the CN. The data traffic is transmitted directly between the mobile network

and CN. As that in other studies of mobile VPN, we mainly consider the security issues when a mobile network moves out of the intranet and wants to access to the resources inside the intranet. Thus, we assume that CN is inside intranet and mobile network moves between intranet and internet. Using Fig. 2 as an example, we consider the cases when the mobile network moves from Home Network to Foreign Network 1, from Foreign Network 1 to Foreign Network 2, then from Foreign Network 2 back to the Home Network.

Fig. 8 illustrates the flows when the mobile network moves from intranet to internet. This may be the case when the mobile network moves from Home Network to Foreign Network 1 as shown in Fig. 2. When a SIP-NVG moves to a foreign network, it must register with the SIP registrar using its new IP address. The registration is similar to that shown in Fig. 5 except that there is no need for SIP header translation. The SIP-NVG then checks whether there are MNs with active sessions inside the mobile network according to the Session Table. The SIP-NVG needs to re-INVITE all CNs to recover

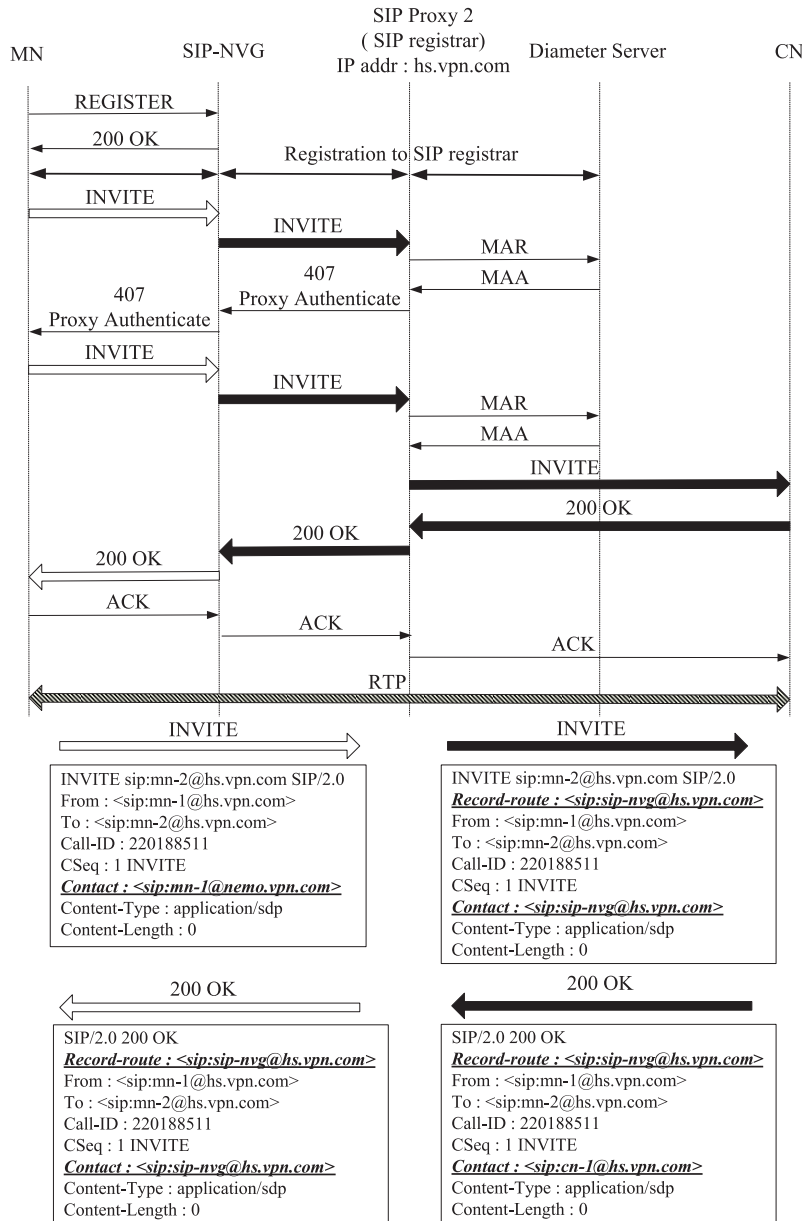


Fig. 6. Message flows and translation of **INVITE** when MN moves to a mobile network which is located inside home network.

all of the ongoing sessions. However, this process may cause substantial signaling messages in wireless links. In order to reduce the signaling overhead, the SIP-NVG combines all contact addresses of CNs into a URI list. The URI list is then conveyed by the SDP embedded in one INVITE message. Fig. 9 shows an example of a URI list in SDP. The re-INVITE contains the new contact address of the SIP-NVG. It is sent to the SIP Proxy 1, which routes the message to the SIP Proxy 2, assuming that SIP Proxy 1 has been informed that SIP Proxy 2 is responsible for the verification of the ongoing sessions.

The proposed SenEMO uses a protocol based on *challenge-response* to authenticate each active session. Username and password are used for verification. When an MN attaches to the mobile network, its SIP-NVG stores such security information to represent the MN. The Diameter command, MAR/MAA, in the Diameter SIP application is used to process authentication messages. To reduce the signaling overhead, the security information for each session is aggregated to one MAR/MAA message. This can be done

easily by setting the reserved bit to “M” in Command Flags within the MAR/MAA’s Diameter header to indicate that there are multiple sessions required to process. Therefore, the SIP Proxy 2 first sends the MAR to request that the Diameter server authenticates and authorizes the sessions. The Diameter server then responds with MAA which contains *challenges* to the SIP Proxy 2. The SIP Proxy 2 uses them to response with 407 Proxy Authenticate to the SIP-NVG. After the SIP-NVG receives the 407 Proxy Authenticate, it sends again the re-INVITE containing authorization which includes username and password for different sessions. The Diameter server verifies the authorization data and informs the SIP Proxy 2 the verification result of each session. If a session is verified successfully, the Diameter Server will generate one pair of TGK and MIKEY for this session. Thereafter, the MAA will have all the verification result codes for each session. To establish TGK and MIKEY message, preshared key, which is one of the most efficient ways to handle key transport, is used. The pairs of the TGK and MIKEY messages

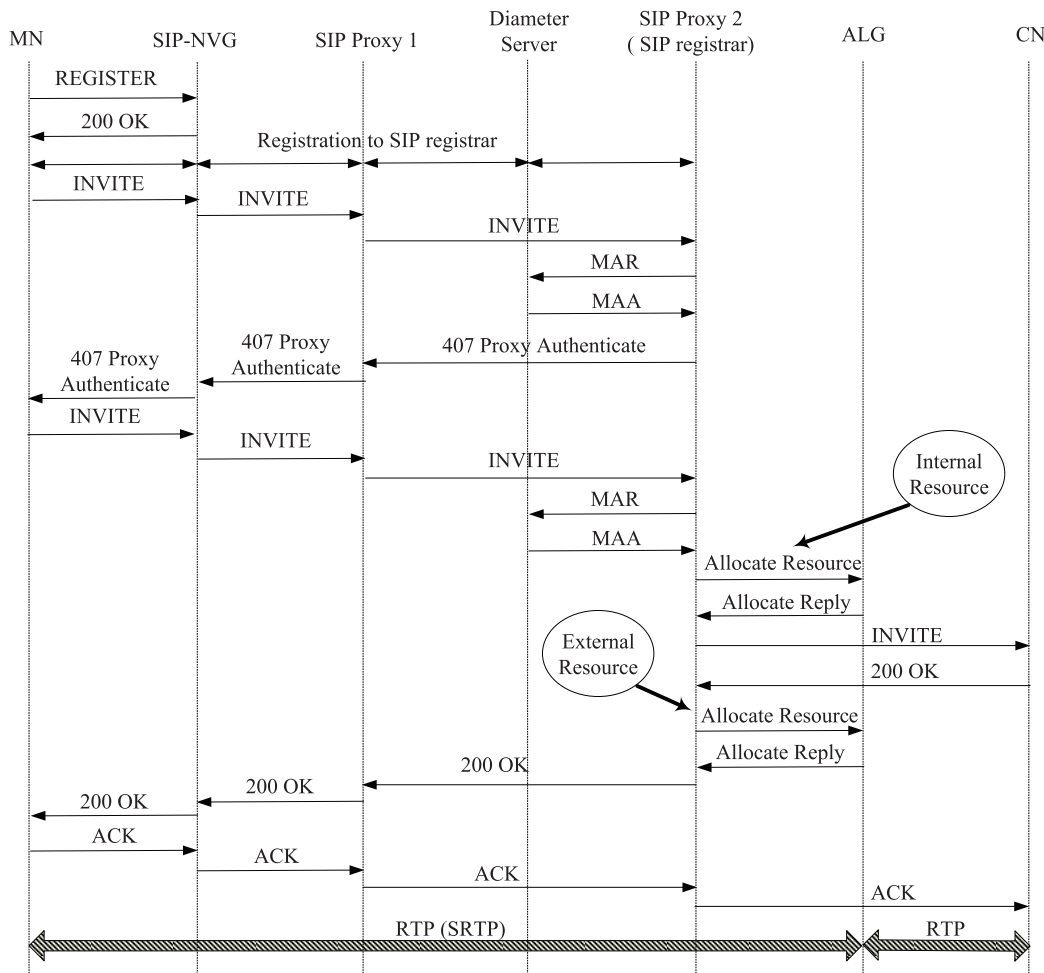


Fig. 7. Message flows when MN moves to a mobile network which is located inside a foreign network.

are transmitted to the SIP Proxy 2 by MAA. We extend the MAA with two more *Attribute Value Pairs (AVPs)*: MIKEY-TGK and MIKEY-MSG. Fig. 10 shows an example of MAA with multiple session information. The extended AVPs are used in MAA only after the session has been authorized by Diameter server and when the communication between MN and CN needs to pass through the ALG.

If the SIP-NVG is granted to access the intranet, the SIP Proxy 2 needs to allocate enough resources to guarantee that the sessions will be protected. First, the SIP Proxy 2 orders the ALG to reserve the internal receiving address and receiving port for each ongoing session. The command from the SIP Proxy 2 may include the Data SA and the TGK for SRTP protection, and the CN's original listening addresses and listening ports. The ALG responds with the reserved addresses and ports. The SIP Proxy 2 inserts the listening addresses and listening ports in the SDP to reproduce INVITE requests based on the URI list. It then routes them to each CN individually. Each CN sends back a *200 OK* if it agrees on the SDP of the re-INVITE. After receiving the *200 OK* from each CN, the SIP Proxy 2 orders again the ALG to reserve the external receiving addresses and receiving ports. The ALG also responds with the reserved external receiving addresses and receiving ports. If all allocated resources are ready, the SIP Proxy 2 replaces the listening addresses and listening ports in the SDP of each *200 OK*. It also inserts the *MIKEY Initiator* message into the SDP to transport TGK.

When the modification is done, the SIP Proxy 2 sends each *200 OK* with the new SDP to the SIP-NVG. Once the SIP-NVG receives each *200 OK*, it will forward the message to the MN. The ALG then will start to function when both internal and external resources have been acquired.

When each MN attaching to the mobile network receives the *200 OK* with the MIKEY Initiator message, it needs to process the MIKEY Initiator message and extract the shared TGK. The MN is then required to send an ACK with SDP which includes the MIKEY responder message. After the MN sends the ACK, it will start all transport sessions. Because the SIP-NVG now is in a foreign network, the session between the MN and the ALG must be reestablished in order to be protected by SRTP. The SIP Proxy 2 exchanges the TGK with the MN on behalf of the CN. As aforementioned discussion, the SIP Proxy 2 transmits the TGK to the ALG which is responsible for converting the protected data streams outside the intranet into unprotected data streams inside the intranet.

Fig. 11 illustrates the flows when the mobile network moves again to another external network. This may be the case when the SIP-NVG moves from Foreign Network 1 to Foreign Network 2 as shown in Fig. 2. The SIP-NVG also needs to update its new location with the Registrar and send re-INVITE to the CNs with active sessions. However, the SIP Proxy 2 will function on behalf of CNs to respond with the *200 OK*. The SIP Proxy 2 only orders the ALG to modify the

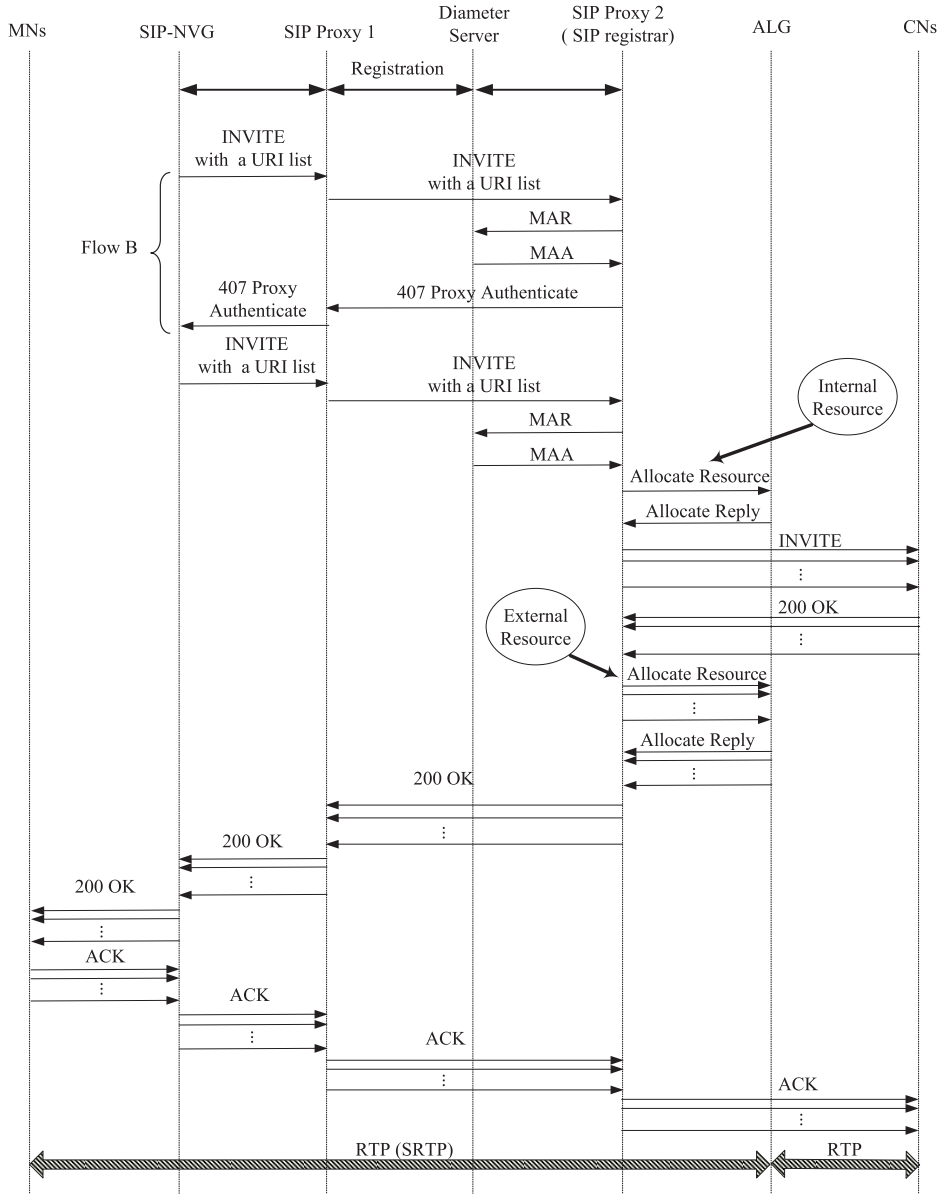


Fig. 8. Message flows when mobile network roams from home network to foreign network.

external listening addresses and ports. If necessary, the SIP Proxy 2 can also re-assign the TGK to the ALG. Therefore, the original data stream will be redirected to the new location of the SIP-NVG. Rest of the discussion is similar to that in Fig. 8.

Fig. 12 depicts the flows when the mobile network moves back to the home network. Because the data traffic is secure inside intranet, the ALG is requested to deallocate both the internal and external resources.

```

v = 0
c = IN IP4 sip-nvg@hs.vpn.com
a = URIlist:<sip:cn-1@hs.vpn.com> port=6600
  ID=132387 src=sip:mn-1@hs.vpn.com
a = URIlist:<sip:cn-2@hs.vpn.com> port=6868
  ID=327623 src=sip:mn-2@hs.vpn.com
    
```

Fig. 9. Example of a URI list with two CNs in SDP.

Because the proposed architecture is based on SIP, the problems inherited from MIP, such as overhead of three tunnels and potential long end-to-end latency, are not problems in our design. Besides, by using SRTP, MIKEY, Diameter SIP Application, and ALG, the proposed architecture can provide VPN services for mobile users effectively. The proposed SeNEMO is particularly useful for real-time applications. However, there might be some security issues in SIP. The following section discusses the security vulnerabilities in SeNEMO.

3.4 Security Vulnerabilities in SeNEMO

Because the proposed SeNEMO is designed based on SIP, the security problems in SeNEMO might be inherited from SIP. The security problems in SIP have been widely studied recently [30], [31], [32], [33], [34]. This section presents the qualitative analysis of security vulnerabilities in the proposed SeNEMO.

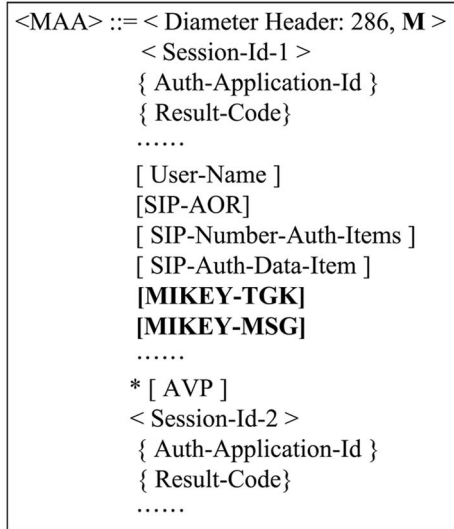


Fig. 10. The extended AVPs of MIKEY-TGK and MIKEY-MSG in MAA.

traffic to get the plaintext or place a spam call. However, in the proposed SeNEMO, the transport of SIP messages can be easily extended to incorporate Transport Layer Security (TLS) [35] so the transmission of SIP messages can be protected.

- Denial of service (DoS): Protecting against DoS attacks is inherently difficult on IP networks, especially for mobile networks such as MIP [36]. Similarly, the open architecture of the Internet may expose any SIP network element, such as Registrar server or Proxy server, to DoS attacks. However, in the proposed SeNEMO, SIP Proxy2/Registrar and Diameter server are located inside the intranet. Therefore, they are much less vulnerable to DoS attacks. Besides, the ALG can block all the messages coming from unknown nodes unless the resources are reserved for legitimate users by SIP Proxy 2. Although SIP Proxy 1 which is on the border may be subject to DoS attacks, a large portion of the effects can be reduced by proper server design and efficient implementation by adequate hardware [31].
- SIP parser attack: The free text format of SIP message could make parsing difficult. Attackers sending a very large messages with unnecessary headers and bodies can exhaust the resource of SIP server. The SeNEMO may suffer from such attack

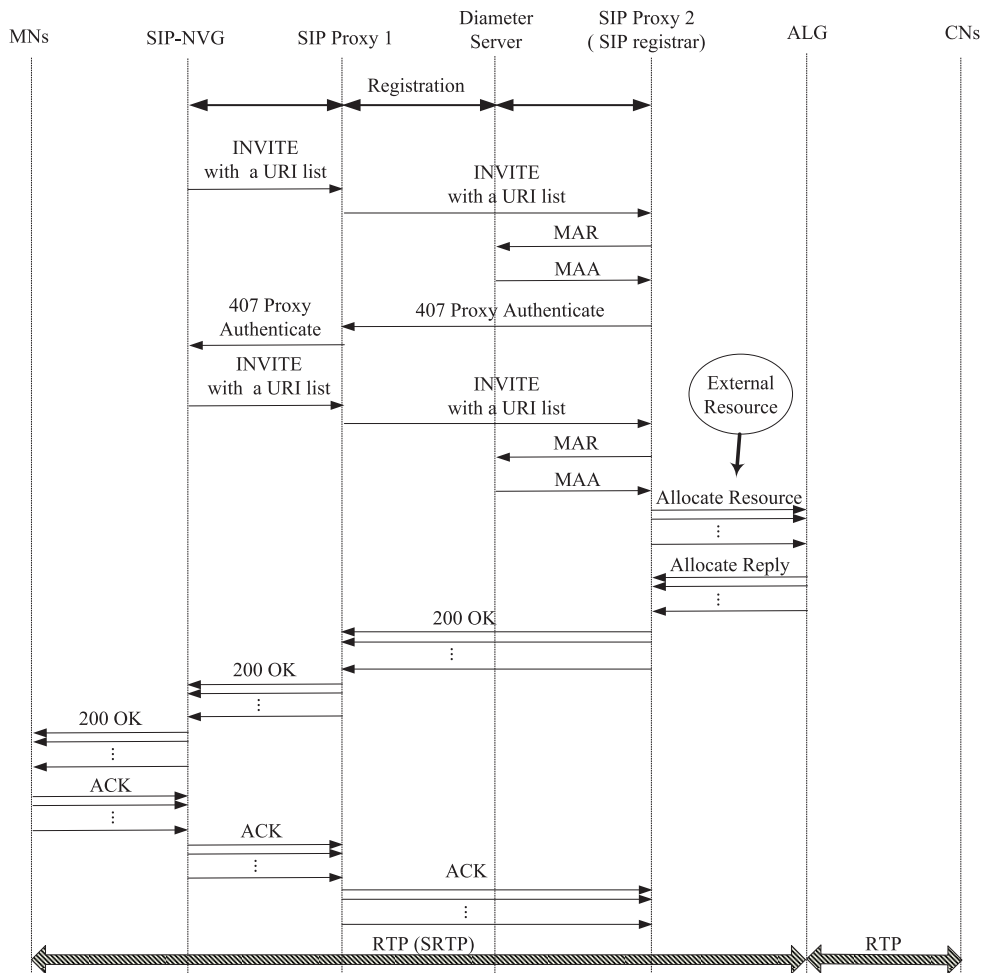


Fig. 11. Message flows when mobile network roams from a foreign network to another foreign network.

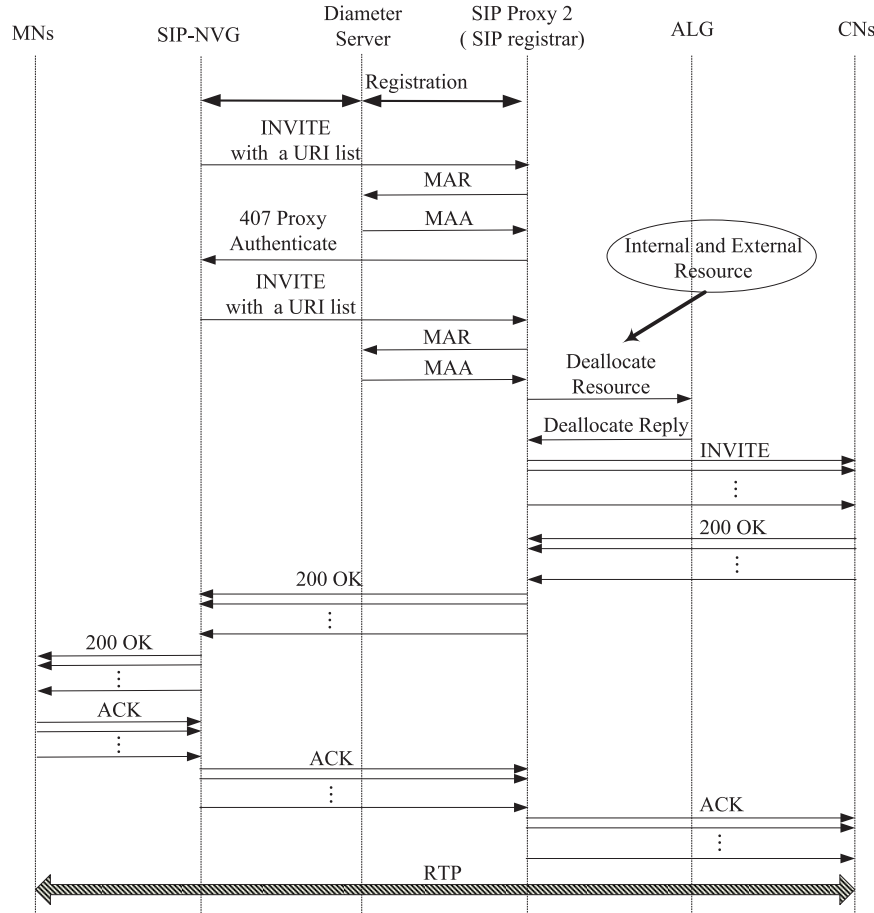


Fig. 12. Message flows when mobile network roams from a foreign network back to home network.

too. To solve this problem, the parser in the SeNEMO can be designed to check message size and discard the one which exceeds the size limit. Also, a practical implementation provided by [33], [34], [37] can be adopted.

- SIP application-level attack: An attacker may send fake BYE, CANCEL, or re-INVITE to terminate a session, cancel an invitation, or redirect a call [30]. In the SeNEMO, such attacks can be prevented by the proposed SIP authentication mechanism and Diameter SIP Application (e.g., *Flow A* in Fig. 5 and *Flow B* in Fig. 8).
- Media security: As mentioned in Section 3.1, the proposed SeNEMO uses MIKEY to transport Data SA and SRTP to secure real-time streaming. Therefore, MNs can securely communicate with CNs even if staying outside the intranet. Please note SRTP provides confidentiality, integrity, and replay protection only for the application data. The UDP and IP headers are not protected as that in IPsec.

The comparison of security vulnerabilities in SeNEMO and IETF MVPN is summarized in Table 2.

4 PERFORMANCE ANALYSIS

In order to support secure communication in VPN, signaling messages carrying security information are sent in the proposed SeNEMO. In addition, signaling messages are sent to maintain session continuity during handoff. To

evaluate the performance of the proposed SeNEMO, it is important to quantify the signaling cost. In this section, we analyze the signaling cost of the proposed SeNEMO. Similar to that in [38], [39], [40], [41], the signaling cost function comprises transmission cost and processing cost. The transmission cost is proportional to the distance between the two network nodes. The processing cost includes the cost to process messages, verify messages, and so on. The proposed SeNEMO has also been implemented. Details can be found in [23].

It is generally agreed that it is difficult to analyze a network mathematically with randomly connected topology. Hence, without loss of generality, we develop an one-dimensional mobility model for handoff. In our proposed SeNEMO, the *Inter-Realm Roaming* of a mobile network includes three types of handoff: 1) handoff from intranet (home network) to a foreign network, 2) handoff from a foreign network to another foreign network, and 3) handoff from a foreign network back to the intranet. They are represented as H_{hf} , H_{ff} , and H_{fh} , respectively. We assume that the network topology is configured as that shown in Fig. 13 in which the mobile network returns to the intranet (home network) after it moves across $N - 1$ foreign networks. Hence, when N is larger, the mobile network travels farther away from its home network before it returns to its home network. For example, the handoff sequence can be shown as $H_{hf}H_{ff}H_{ff}\dots H_{ff}H_{fh}$. Table 3 lists the parameters used in the analysis.

TABLE 2
Comparison of Security Vulnerabilities in SeNEMO and IETF MVPN

	SeNEMO	IETF MVPN
Key management	MIKEY (especially designed for real-time applications running over SRTP)	IKE (more heavy for real-time applications)
Data transport	SRTP (supporting confidentiality and integrity with lower cost for bandwidth; headers are not protected)	IPsec (supporting confidentiality and integrity; headers are protected)
AAA	Diameter SIP Application	Diameter MIPv4 Application
Network components in internet	N/A	Where to put x-HA and how to trust x-HA?
DoS	Less vulnerable to DoS attacks (all nodes are located inside intranet)	More vulnerable to DoS attacks (x-HA is located in internet)

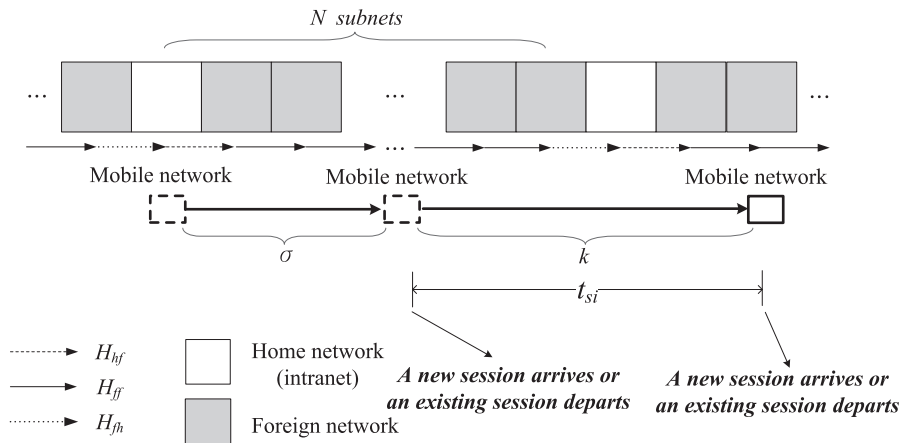


Fig. 13. Network topology for analysis.

For a mobile network, let $f_m(t)$ be a general density function for the network residence time t_M in a subnet. Let $E[t_M] = 1/\gamma$. Its Laplace transform is written as:

$$f_m^*(s) = \int_{t=0}^{\infty} e^{-st} f_m(t) dt.$$

We assume that the arrival of SIP sessions to the mobile network follows Poisson process with arrival rate λ . The service time of a session is exponentially distributed with mean $1/\mu$. Considering the diagram in Fig. 13, similar to that in [42], we define $\alpha_i(k)$ as the probability when a mobile network moves across k subnets between two events while there are i ongoing sessions in the mobile network. The event here refers to that a new session arrives or an ongoing session departs from the mobile network. We denote t_{s_i} as the interval between two consecutive events. During t_{s_i} there are i ongoing sessions in the mobile network. Based on the property of sums of two independent Poisson process, t_{s_i} can be considered as the inter arrival time of a new Poisson process. Therefore,

$$E[t_{s_i}] = \frac{1}{\pi_i} = \begin{cases} \frac{1}{\lambda + i\mu}, & 0 \leq i < c \\ \frac{1}{i\mu}, & i = c, \end{cases} \quad (1)$$

where c is the maximum number of ongoing sessions allowed in a mobile network.

According to above assumption, we have:

$$\alpha_i(k) = \begin{cases} 1 - \frac{(1 - f_m^*(\pi_i))\gamma}{\pi_i}, & k = 0, \\ \frac{\gamma}{\pi_i} [1 - f_m^*(\pi_i)]^2 [f_m^*(\pi_i)]^{k-1}, & k > 0. \end{cases} \quad (2)$$

For simplicity, we denote $g_i = f_m^*(\pi_i)$ in the rest of the paper. Let $k = jN + q$ and $0 \leq q < N$. Then,

$$\alpha_i(jN + q) = \frac{\gamma(1 - g_i)^2}{\pi_i g_i} (g_i^N)^j g_i^q = yz^j x^q, \quad (3)$$

where $y = \frac{\gamma(1 - g_i)^2}{\pi_i g_i}$, $z = g_i^N$, $x = g_i$.

For demonstration purpose, we assume that the network residence time follows Gamma distribution. The Laplace transform of a Gamma random variable is expressed as:

TABLE 3
List of Parameters Used in Analysis

N	Number of networks a mobile network visits before it goes back to intranet.
λ	Session arrival rate for a mobile network
$1/\mu$	Average session service time
$1/\gamma$	Average network residence time
c	Maximum number of ongoing sessions in a mobile network
δ_σ	Probability density function of σ

$$f_m^*(s) = \left(\frac{\gamma\beta}{s + \gamma\beta} \right)^\beta. \quad (4)$$

Hence, we obtain:

$$g_i = f_m^*(\pi_i) = \left(\frac{\gamma\beta}{\pi_i + \gamma\beta} \right)^\beta. \quad (5)$$

In the proposed SeNEMO, when a mobile network moves across networks, it needs to perform registration with the SIP Registrar to update its location. It also needs to send re-INVITE messages to CNs if there are ongoing sessions with the MNs in the mobile network. Hence, the cost comprises two parts: the registration cost for SIP-NVG and the re-INVITE cost for maintaining session continuity. The registration cost is independent of the number of ongoing sessions in the mobile network because the SIP-NVG can register with the SIP Registrar on behalf of the whole mobile network. On the other hand, the re-INVITE cost depends on the number of ongoing sessions in the mobile network. The cost increases when the number of ongoing sessions increases. However, because we design a URI list embedded in one re-INVITE message, the cost for the re-INVITE before really sending re-INVITE message to each individual CN is nearly constant regardless of how many ongoing sessions in the mobile network. It can be seen in Figs. 8, 11, and 12. We define the following parameters:

- S_{hf}^i : Average handoff cost when a mobile network moves from home network to foreign network with i ongoing sessions.
- S_{ff}^i : Average handoff cost when a mobile network moves from a foreign network to another foreign network with i ongoing sessions.
- S_{fh}^i : Average handoff cost when a mobile network moves from foreign network to home network with i ongoing sessions.
- R_h : Average registration cost of a mobile network (sent by SIP-NVG) when the mobile network enters its home network.
- R_f : Average registration cost of a mobile network (sent by SIP-NVG) when the mobile network enters a foreign network.
- L_{hf} : Average cost for the first part of re-INVITE when a mobile network moves from its home network to a foreign network.
- L_{ff} : Average cost for the first part of re-INVITE when a mobile network moves from a foreign network to another foreign network.
- L_{fh} : Average cost for the first part of re-INVITE when a mobile network moves from a foreign network to its home network.
- I_{hf} : Average cost for the second part of re-INVITE of a session when a mobile network moves from its home network to a foreign network.
- I_{ff} : Average cost for the second part of re-INVITE of a session when a mobile network moves from a foreign network to another foreign network.
- I_{fh} : Average cost for the second part of re-INVITE of a session when a mobile network moves from a foreign network to its home network.

Therefore, we can denote the signaling cost for handoff as:

$$\begin{aligned} S_{hf}^i &= R_f + L_{hf} + iI_{hf}, \\ S_{ff}^i &= R_f + L_{ff} + iI_{ff}, \\ S_{fh}^i &= R_h + L_{fh} + iI_{fh}. \end{aligned} \quad (6)$$

During t_{S_i} , we assume that the mobile network crosses k subnets as that shown in Fig. 13. We also denote σ as the number of subnets the mobile network moved across from the time it visited the intranet until the time the last event occurred. When $0 < \sigma < N$, the total signaling cost in the mobile network during t_{S_i} can be derived as:

$$\begin{aligned} C_i(N, \pi_i, \gamma, \sigma) &= \sum_{k=0}^{\infty} \left\{ S_{hf}^i \left\lfloor \frac{k + \sigma - 1}{N} \right\rfloor + S_{fh}^i \left\lfloor \frac{k + \sigma}{N} \right\rfloor \right. \\ &\quad \left. + S_{ff}^i \left(k - \left\lfloor \frac{k + \sigma - 1}{N} \right\rfloor - \left\lfloor \frac{k + \sigma}{N} \right\rfloor \right) \right\} \alpha_i(k) \\ &= S_{ff}^i \frac{\gamma}{\pi_i} + \frac{(S_{hf}^i g_i - S_{ff}^i g_i + S_{fh}^i - S_{ff}^i)(1 - g_i) g_i^{N-1} \gamma}{(1 - g_i^N) g_i^\sigma \pi_i}. \end{aligned} \quad (7)$$

If $\sigma = 0$, that is, an event occurs when the mobile network is located inside the intranet, the total signaling cost during t_{S_i} can be derived as:

$$\begin{aligned} C_i(N, \pi_i, \gamma, 0) &= \sum_{k=0}^{\infty} \left\{ S_{hf}^i \left\lfloor \frac{k}{N} \right\rfloor + S_{fh}^i \left\lfloor \frac{k}{N} \right\rfloor + S_{ff}^i \left(k - \left\lfloor \frac{k}{N} \right\rfloor - \left\lfloor \frac{k}{N} \right\rfloor \right) \right\} \alpha_i(k) \\ &= S_{ff}^i \frac{\gamma}{\pi_i} + \frac{(S_{hf}^i - S_{ff}^i + (S_{fh}^i - S_{ff}^i) g_i^{N-1})(1 - g_i) \gamma}{(1 - g_i^N) \pi_i}. \end{aligned} \quad (8)$$

Moreover, if σ is uniformly distributed with $\delta_\sigma = 1/N$, we can rewrite (8) as:

$$\begin{aligned} C_{i_uniform}(N, \pi_i, \gamma) &= \frac{1}{N} \sum_{\sigma=0}^{N-1} C_i(N, \pi_i, \gamma, \sigma) \\ &= S_{ff}^i \frac{\gamma}{\pi_i} + \frac{1}{N} \left\{ \frac{(S_{hf}^i - S_{ff}^i + (S_{fh}^i - S_{ff}^i) g_i^{N-1})(1 - g_i) \gamma}{(1 - g_i^N) \pi_i} \right. \\ &\quad \left. + \sum_{\sigma=1}^{N-1} \frac{(S_{hf}^i g_i - S_{ff}^i g_i + S_{fh}^i - S_{ff}^i)(1 - g_i) g_i^{N-1} \gamma}{(1 - g_i^N) g_i^\sigma \pi_i} \right\} \\ &= \frac{\gamma}{\pi_i} \left(S_{ff}^i + \frac{S_{hf}^i - 2S_{ff}^i + S_{fh}^i}{N} \right). \end{aligned} \quad (9)$$

If most of time a mobile network is located inside intranet or a foreign network near the intranet, we can assume that σ is linearly distributed. The p.d.f of σ is expressed as:

$$\delta_\sigma = \begin{cases} \left(\frac{N+1}{2} - \sigma \right) \left(\frac{2}{N+1} \right)^2, & 0 \leq \sigma \leq \frac{N-1}{2} \\ \left(\sigma - \frac{N-1}{2} \right) \left(\frac{2}{N+1} \right)^2, & \frac{N-1}{2} < \sigma \leq N-1. \end{cases} \quad (10)$$

We then can get:

$$\begin{aligned}
C_{i_linear}(N, \pi_i, \gamma) &= \sum_{\sigma=0}^{N-1} \delta_{\sigma} C_i(N, \pi_i, \gamma, \sigma) \\
&= S_{ff}^i \frac{\gamma}{\pi_i} + \frac{\gamma(1-g_i)}{\pi_i(1-g_i^N)} \left\{ \frac{2(S_{hf}^i - S_{ff}^i + (S_{fh}^i - S_{ff}^i)g_i^{N-1})}{N+1} \right. \\
&\quad \left. + \frac{4(S_{hf}^i g_i - S_{ff}^i g_i + S_{fh}^i - S_{ff}^i)g_i^{N-1}}{(N+1)^2} \right. \\
&\quad \left. \left(\sum_{\sigma=1}^{(N-1)/2} \frac{1}{g_i^{\sigma}} \left(\frac{N+1}{2} - \sigma \right) + \sum_{\sigma=(N+1)/2}^{N-1} \frac{1}{g_i^{\sigma}} \left(\sigma - \frac{N-1}{2} \right) \right) \right\}. \tag{11}
\end{aligned}$$

If σ is exponentially distributed, the p.d.f of σ is:

$$\delta_{\sigma} = \begin{cases} e^{-\sigma} \left(\frac{1 - 2e^{-(N+1)/2} + e^{-1}}{1 - e^{-1}} \right), & 0 \leq \sigma \leq \frac{N-1}{2} \\ e^{-(N-\sigma)} \left(\frac{1 - 2e^{-(N+1)/2} + e^{-1}}{1 - e^{-1}} \right), & \frac{N-1}{2} < \sigma \leq N-1. \end{cases} \tag{12}$$

Therefore,

$$\begin{aligned}
C_{i_exponential}(N, \pi_i, \gamma) &= \sum_{\sigma=0}^{N-1} \delta_{\sigma} C_i(N, \pi_i, \gamma, \sigma) = S_{ff}^i \frac{\gamma}{\pi_i} + \frac{\gamma(1-g_i)}{\pi_i(1-g_i^N)} \\
&\quad \left\{ \frac{(1-e^{-1})(S_{hf}^i - S_{ff}^i + (S_{fh}^i - S_{ff}^i)g_i^{N-1})}{1 - 2e^{-(N+1)/2} + e^{-1}} \right. \\
&\quad \left. + \frac{(1-e^{-1})(S_{hf}^i g_i - S_{ff}^i g_i + S_{fh}^i - S_{ff}^i)g_i^{N-1}}{1 - 2e^{-(N+1)/2} + e^{-1}} \right. \\
&\quad \left. \left(\sum_{\sigma=1}^{(N-1)/2} \frac{e^{-\sigma}}{g_i^{\sigma}} + \sum_{\sigma=(N+1)/2}^{N-1} \frac{e^{-(N-\sigma)}}{g_i^{\sigma}} \right) \right\}. \tag{13}
\end{aligned}$$

As aforementioned discussion, the arrival of sessions to a mobile network follows Poisson process and the session service time is exponentially distributed. In addition, there is a limit c for the maximum number of ongoing sessions allowed in the mobile network. Therefore, we can model the number of ongoing sessions in a mobile network as $M/M/c/c$ queuing system. The steady state probability that there are i ongoing sessions in the mobile network is given by [43]:

$$P_i = \frac{\lambda^i}{i! \mu^i} \left(\sum_{j=0}^c \frac{\lambda^j}{j! \mu^j} \right)^{-1}. \tag{14}$$

As a result, the average handoff-signaling cost per unit time can be derived as:

$$\sum_{i=0}^c C_{i_d} P_i \pi_i. \tag{15}$$

where δ can be *uniform*, *linear*, or *exponential* distribution. The variables of π_i, P_i can be obtained from (1) and (14).

To evaluate the performance of the proposed SeNEMO, we define the following parameters:

- a_x : The processing cost for SIP registration at Node x .
- b_x : The processing cost for SIP INVITE message at Node x .

- $A_{x,y}$: The transmission cost of SIP registration between Node x and Node y .
- $B_{x,y}$: The transmission cost of SIP INVITE message between Node x and Node y .
- U : The total cost for SIP Proxy 1 to process and transmit UAR/UAA messages to Diameter Server.
- M : The total cost for SIP Proxy 2 to process and transmit MAR/MAA messages to Diameter Server.

where x and y can be mn, nvg, pro, reg, alg or cn which denote MN, SIP-NVG, SIP Proxy 1, SIP Proxy 2 (SIP Registrar), ALG, and CN, respectively.

According to the signaling message flows described in Section 3, the above cost can be calculated as:

$$\begin{aligned}
R_h &= a_{nvg} + 2a_{reg} + 4A_{nvg,reg} + 2M, \\
R_f &= a_{nvg} + 4a_{pro} + 2a_{reg} + 4A_{nvg,pro} + 4A_{pro,reg} + 2U + 2M, \\
L_{hf} &= b_{nvg} + 3b_{pro} + 3b_{reg} + b_{alg} + 3B_{nvg,pro} + 3B_{pro,reg} \\
&\quad + 2B_{reg,alg} + 2M, \\
L_{ff} &= b_{nvg} + 3b_{pro} + 2b_{reg} + 3B_{nvg,pro} + 3B_{pro,reg} + 2M, \\
L_{fh} &= b_{nvg} + 3b_{reg} + b_{alg} + 3B_{nvg,reg} + 2B_{reg,alg} + 2M, \\
I_{hf} &= b_{mn} + 2b_{nvg} + 2b_{pro} + 3b_{reg} + b_{alg} + b_{cn} + 2B_{mn,nvg} \\
&\quad + 2B_{nvg,pro} + 2B_{pro,reg} + 3B_{reg,cn} + 2B_{reg,alg}, \\
I_{ff} &= b_{mn} + 2b_{nvg} + 2b_{pro} + b_{reg} + b_{alg} + 2B_{mn,nvg} + 2B_{nvg,pro} \\
&\quad + 2B_{pro,reg} + 2B_{reg,alg}, \\
I_{fh} &= b_{mn} + 2b_{nvg} + 2b_{reg} + b_{cn} + 2B_{mn,nvg} + 2B_{nvg,reg} + 3B_{reg,cn}.
\end{aligned}$$

To compare the signaling cost with IETF MVPN, we assume the Diameter MIPv4 Application [44] is used to authenticate the x-MIP. Also, we assume MN is in collocated mode. The analysis of FA mode is not presented because it has almost same results. The subscripts mn, mg, vpn, xha , and iha refer to MN, Mobile Gateway, IPsec-based VPN gateway, x-HA and i-HA, respectively. Also, the following parameters are defined:

- M_x : The x-MIP registration cost.
- M_{i-o} : The i-MIP registration cost when MN is located outside intranet.
- M_{i-i} : The i-MIP registration cost when MN is located inside intranet.
- T_{est} : The establishment cost of IPsec tunnel.
- T_{ter} : The termination cost of IPsec tunnel.
- d_x : The processing cost for MIP registration at Node x .
- e_x : The processing cost for IPsec message at Node x .
- $W_{x,y}$: The transmission cost of MIP registration between Node x and Node y .
- $Z_{x,y}$: The transmission cost of IPsec message between Node x and Node y .
- H : The total cost for x-HA to process and transmit HAR/HAA and AMR/AMA messages to AAAF and AAAH.

Based on the signaling message flows shown in Fig. 14, the above cost can be calculated as:

$$\begin{aligned}
M_x &= 2d_{mg} + 2d_{xha} + 2W_{mn,mg} + 2W_{mg,xha} + 2H, \\
M_{i-o} &= 2d_{mg} + 2d_{xha} + 2d_{vpn} + d_{iha} + 2W_{mn,mg} + 2W_{mg,xha} \\
&\quad + 2W_{xha,vpn} + 2W_{vpn,iha} + 2H, \\
M_{i-i} &= d_{iha} + 2W_{mn,mg} + 2W_{mg,iha}, \\
T_{est} &= 2e_{mn} + 6e_{mg} + 6e_{xha} + 3e_{vpn} + 6Z_{mn,mg} + 6Z_{mg,xha} \\
&\quad + 6Z_{xha,vpn}, \\
T_{ter} &= Z_{mn,mg} + Z_{mg,vpn} + e_{vpn}.
\end{aligned}$$

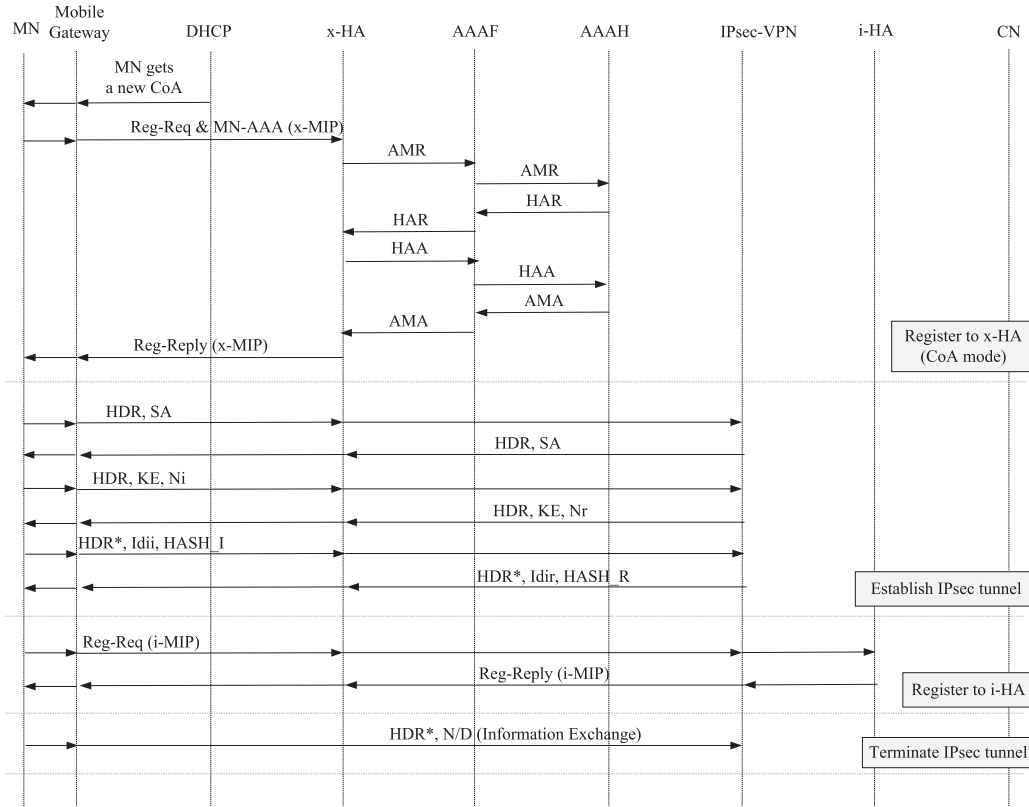


Fig. 14. Signaling flows of IETF MVPN.

The handoff cost of IETF MVPN when a mobile network moves between different networks is derived as:

$$\begin{aligned} D_{hf} &= M_x + M_{i-o} + T_{est}, \\ D_{ff} &= M_x, \\ D_{fh} &= M_{i-i} + T_{ter}. \end{aligned} \quad (16)$$

5 NUMERICAL RESULTS

This section provides the numerical results for the analysis presented in Section 4. The analysis is validated by extensive simulations by using ns-2 [45]. As that discussed in Section 4, the signaling cost function consists of transmission cost and processing cost. We assume that the transmission cost is proportional to the distance between the source and destination nodes, and the processing cost includes the processing and verifying SIP messages [38], [39], [40], [41]. Besides, the transmission cost of wireless link is higher than that of wireline. To illustrate the performance, the values reasonably chosen for the parameters are listed in Table 4. Besides, to compare with the signaling cost of the IETF MVPN, we assume the x-HA is optimally colocated with the VPN gateway and AAAF, and the i-HA is located in the same position of the SIP Proxy 2/Registrar. The AAAH in IETF MVPN is located in the same position of the Diameter server in the SeNEMO. The objective of the analysis is to quantify the performance. Choosing other values will not change the conclusion drawn from the analysis.

In the proposed SeNEMO, one of the major objectives is to reduce the signaling cost for handoff while still support VPN. In our proposed architecture, the SIP-NVG manages the mobility of the entire mobile network. Therefore, once the

mobile network moves to a new subnet, the SIP-NVG can register with the SIP Registrar for the whole mobile network. On the other hand, if there is no SIP-NVG, every MN in the same mobile network needs to update its location individually. Thus, more signaling cost will be incurred. We can redefine (6) for the cost when there is no SIP-NVG:

TABLE 4
Parameters for Performance Evaluation

Parameter	Value	Parameter	Value
a_{nvg}	25.0	$A_{nvg,reg}$	5.0
a_{pro}	5.0	$A_{nvg,pro}$	10.0
a_{reg}	25.0	$A_{pro,reg}$	0.1
b_{mn}	5.0	$B_{mn,nvg}$	1.0
b_{nvg}	25.0	$B_{nvg,reg}$	5.0
b_{pro}	5.0	$B_{nvg,pro}$	10.0
b_{reg}	40.0	$B_{pro,reg}$	0.1
b_{alg}	50.0	$B_{reg,alg}$	0.1
b_{cn}	10.0	$B_{reg,cn}$	0.5
U	30.0	M	40.0
d_{mg}	10.0	$W_{mn,mg}$	1.0
d_{xha}	25.0	$W_{mg,xha}$	10.0
d_{vpn}	40.0	$W_{xha,vpn}$	1.0
d_{iha}	25.0	$W_{vpn,iha}$	0.1
H	80.0	$W_{mg,iha}$	5.0
e_{mn}	5.0	$Z_{mn,mg}$	1.0
e_{mg}	10.0	$Z_{mg,xha}$	10.0
e_{xha}	25.0	$Z_{xha,vpn}$	1.0
e_{vpn}	40.0	$Z_{mg,vpn}$	11.0

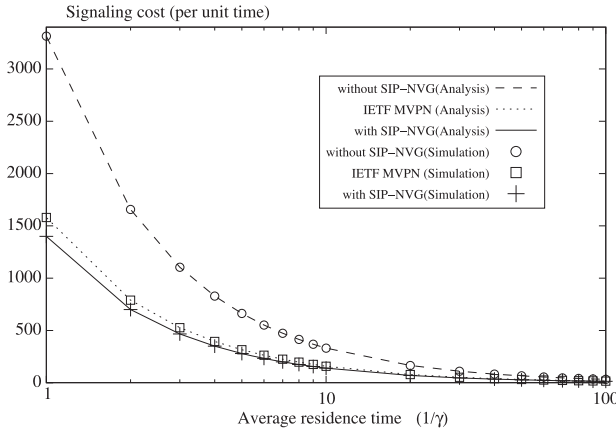


Fig. 15. Comparison of signaling cost with different residence time.

$$\begin{aligned}
 S_{hf}^i &= mR_f + iL_{hf} + iI_{hf}, \\
 S_{ff}^i &= mR_f + iL_{fh} + iI_{ff}, \\
 S_{fh}^i &= mR_h + iL_{hh} + iI_{fh}.
 \end{aligned} \quad (17)$$

where we assume m is the number of MNs attaching to the mobile network.

In Fig. 15, the average signaling costs with and without SIP-NVG are compared with IETF MVPN. As defined above, m is the number of MNs attaching to the mobile network. In addition, we assume $N = 7$, $m = 5$, $c = 10$, and $\rho = \lambda/\mu = 5$. In the SIP-based protocol, the SIP re-INVITES and SIPs registration need to be performed during each handoff. Therefore, the signaling cost for SIP-based protocol might be higher than that in MIP. However, Fig. 15 shows that SeNEMO with SIP-NVG has smaller signaling cost for handoff than that in IETF MVPN. This is because IETF MVPN requires certain time to establish the three tunnels. In addition, IETF MVPN is designed only for single node mobility. When comparing with the mobile network without SIP-NVG, as expected, the proposed SeNEMO reduces handoff signaling cost significantly. We can also see that when the average network residence time increases, that is the mobile network has relative low mobility, the signaling cost for handoff decreases.

Fig. 16 demonstrates the average signaling cost for handoff versus ρ . The parameters are set as that in Fig. 15 except $\gamma = 0.1$. Similar to Fig. 15, the proposed SeNEMO has less signaling cost for handoff than that without SIP-NVG. On the other hand, IETF MVPN has higher or less signaling cost than SeNEMO depending on ρ . This is because the signaling cost for handoff in MIP is independent of session number. MIP only deals with mobility management at IP layer. Therefore, the signaling cost for handoff keeps constant regardless of the session capacity and ρ , which represents the number of sessions in the mobile network. However, in the analysis for IETF MVPN, we consider the lower bound for IETF MVPN. The upper bound is difficult to obtain because it depends on many other factors. In Fig. 16, the costs for periodical location update that an MN sends to x-HA and i-HA are not considered. Also, we assume the x-HA is placed in an optimal position. When these factors are included, the cost for IETF MVPN will be significantly higher.

Although Fig. 16 shows that the signal cost for SeNEMO may be higher than that in IETF MVPN, the packet deliver

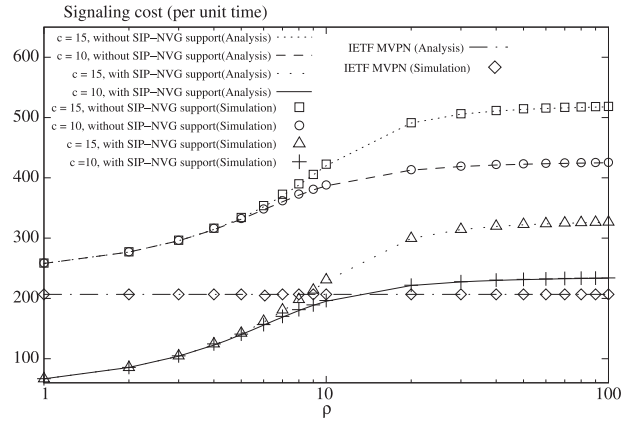


Fig. 16. Comparison of signaling cost with and without SIP-NVG.

cost in SeNEMO is much lower than that in IETF MVPN. This is because in SeNEMO, there is no issue such as triangular routing and three tunnels. We also see that when ρ increases, the average cost for SIP-based solutions increases too. The reason is that with more ongoing sessions, more re-INVITES are needed to maintain session continuity. Besides, when ρ is larger than 20, the costs of all techniques presented in Fig. 16 almost remain constant. This is because when ρ approaches to 20, the number of ongoing sessions of each technique reaches the maximum number allowed in the mobile network. Comparing $c = 10$ and $c = 15$ in the SIP-based techniques, more sessions exist in the mobile network when $c = 15$. Hence more signaling cost is produced for re-INVITE.

In Section 4, linear and exponential distributions are used if most of time a mobile network is located inside intranet or a foreign network near the intranet. Fig. 17 shows the performance results of different distributions when SeNEMO is used. We set the parameters as $\gamma = 0.1$, $N = 7$ and $c = 10$. One can see that the cost for exponential distribution is highest. This is because the probability that the mobile network is located inside the intranet is highest for exponential distribution. Because moving from home network to foreign network causes higher signaling cost than other types of handoff, revisiting the home network frequently will result in more signaling cost. Due to the space limit, the results of IETF MVPN and the one without SIP-NVG are not shown. However, they reach the same conclusion discussed here.

Furthermore, we consider the effect of the variance of mobility pattern. We assume the average residence time is gamma distributed [46], [47]. Therefore, its variance is:

$$\text{Var} = \frac{1}{\beta\gamma^2}. \quad (18)$$

Figs. 18 and 19 illustrate the *signaling cost ratio* with different variances. Here, we define the *signaling cost ratio* as the ratio of the average signaling cost of the proposed SeNEMO over the average signaling cost of the architecture without SIP-NVG. Figs. 18 and 19 show the results for linear distribution and exponential distribution of δ_{σ} , respectively. According to (18), smaller β means larger variance. Because the proposed SeNEMO can reduce the signaling cost for handoff, the signaling cost ratio is always below 1.0 for any session arrival rate. In addition, it appears that the variance does not significantly affect the performance of the

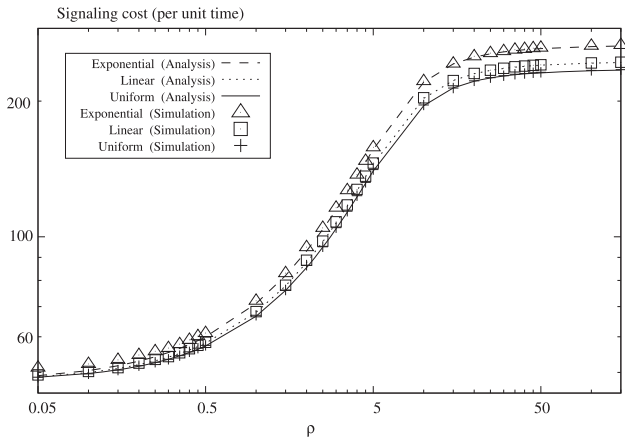


Fig. 17. Comparison of signaling cost with different distribution (δ_o).

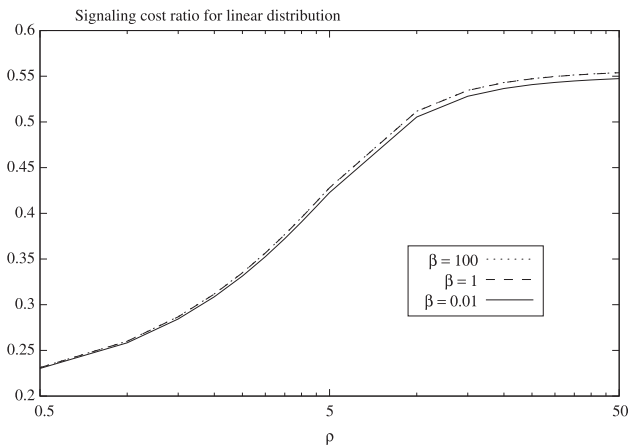


Fig. 18. Effect of the variance of residence time (linear distribution).

proposed SeNEMO. Comparing Figs. 18 and 19, both of them have similar results.

6 CONCLUSIONS

Although IETF has proposed a mobile VPN architecture, it is designed for the movement of a signal node. In addition, the IETF MVPN has large overhead for transmitting real-time packets because it requires one IPsec tunnel and two MIP tunnels. On the other hand, there is no efficient way to support mobile VPN in NEMO although NEMO support network mobility. In this paper, we present the design and analysis of SeNEMO, which integrates NEMO and VPN. The proposed SeNEMO is based on SIP, which makes it particularly suitable for real-time services. Although SIP-based mobility management is easy to support route optimization, by adopting SIP into NEMO, it may increase signaling cost during handoff due to sending many re-INVITE messages for ongoing sessions. In our proposed SeNEMO, a URI list is used to inform SIP Proxy Server instead of sending the information for each MN individually. Thus, signaling cost is reduced. Various IETF protocols have been adopted in SeNEMO. The SIP Proxy Server and the Diameter Server are responsible for authentication and authorization. By following the MIDCOM architecture, the ALG accepts the commands from SIP Proxy Server to process the security information for data transmission. For key management, we adopt MIKEY to exchange the shared

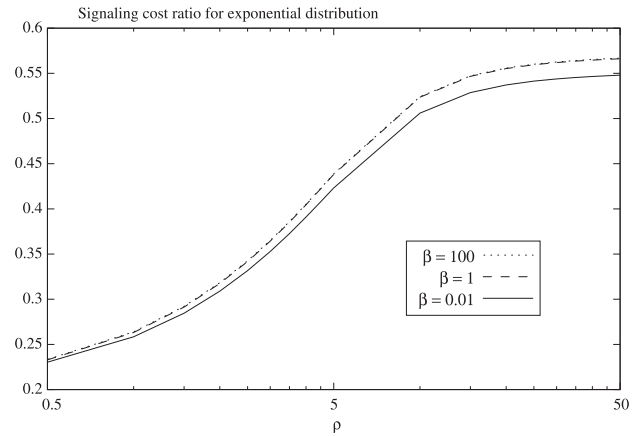


Fig. 19. Effect of the variance of residence time (exponential distribution).

TGK, that is used to protect data transport. The Diameter Server exchanges the TGK with the MN on behalf of CN. The SIP Proxy Server will transmit TGK to ALG. The ALG is responsible for switching and relaying the protected and unprotected data. Therefore, unauthorized data cannot pass through ALG into the intranet. In addition to analyzing the security vulnerabilities, we also propose analytical models to compare the performance of the proposed SeNEMO with IETF MVPN.

ACKNOWLEDGMENTS

This work was sponsored in part by the National Science Council (NSC) under Grant 98-2221-E-009-187-MY3.

REFERENCES

- [1] V. Schena and G. Losquadro, "FIFTH Project Solutions Demonstrating New Satellite Broadband Communication System for High Speed Train," *Proc. IEEE Vehicular Technology Conf.*, pp. 2831-2835, May 2004.
- [2] "WirelessCabin Project," <http://www.wirelesscabin.com>, 2011.
- [3] V. Devarapalli, R. Wakikawa, A. Petrescu, and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol," IETF RFC 3963, Jan. 2005.
- [4] C.E. Perkins, "IP Mobility Support for IPv4," IETF RFC 3344, 2002.
- [5] S. Vaarala and E. Klovning, "Mobile IPv4 Traversal Across IPsec-Based VPN Gateways," IETF RFC 5265, June 2008.
- [6] S.-C. Huang, Z.-H. Liu, and J.-C. Chen, "SIP-Based Mobile VPN for Real-Time Applications," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC '05)*, pp. 2318-2323, Mar. 2005.
- [7] Z.-H. Liu, J.-C. Chen, and T.-C. Chen, "Design and Analysis of SIP-Based Mobile VPN for Real-Time Applications," *IEEE Trans. Wireless Comm.*, vol. 8, no. 11, pp. 5650-5661, Nov. 2009.
- [8] S. Kent and R. Atkinson, "Security Architecture for the Internet Protocol," IETF RFC 2401, Nov. 1998.
- [9] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiation Protocol," IETF RFC 3261, June 2002.
- [10] D. Harkins and D. Carrel, "The Internet Key Exchange (IKE)," IETF RFC 2409, Nov. 1998.
- [11] J.-C. Chen, Y.-W. Liu, and L.-W. Lin, "Mobile Virtual Private Networks with Dynamic MIP Home Agent Assignment," *Wireless Comm. and Mobile Computing*, vol. 6, no. 5, pp. 601-616, Aug. 2006.
- [12] J.-C. Chen, J.-C. Liang, S.-T. Wang, S.-Y. Pan, Y.-S. Chen, and Y.-Y. Chen, "Fast Handoff in Mobile Virtual Private Networks," *Proc. IEEE Int'l Symp. World of Wireless Mobile and Multimedia Networks (WoWMoM '06)*, pp. 548-552, June 2006.
- [13] D. Collins, *Carrier Grade Voice over IP*, second ed. McGraw-Hill, Sept. 2002.

- [14] H.-J. Lim, D.-Y. Lee, and T.-M. Chung, "Comparative Analysis of IPv6 VPN Transition in NEMO Environments," *Proc. Int'l Conf. Computational Science and Its Applications*, pp. 486-496, May 2006.
- [15] T.K. Tan and A. Samsudin, "Efficient NEMO Security Management via CAP-KI," *Proc. IEEE Int'l Conf. Telecomm. and Malaysia Int'l Conf. Comm. (ICT-MICC '07)*, pp. 140-144, May 2007.
- [16] A. Dutta, F. Vakil, J.-C. Chen, M. Tauli, S. Baba, N. Nakajima, and H. Schulzrinne, "Application Layer Mobility Management Scheme for Wireless Internet," *Proc. IEEE Int'l Conf. Third Generation Wireless and beyond (3G Wireless)*, pp. 379-385, May 2001.
- [17] D. Vali, S. Paskalis, A. Kaloxylos, and L. Merakos, "An Efficient Micro-Mobility Solution for SIP Networks," *Proc. IEEE GLOBECOM*, pp. 3088-3092, Dec. 2003.
- [18] S. Pack, X. Shen, J.W. Mark, and J. Pan, "Mobility Management in Mobile Hotspots with Heterogeneous Multihop Wireless Links," *IEEE Comm. Magazine*, vol. 45, no. 9, pp. 106-112, Sept. 2007.
- [19] C.-M. Huang, C.-H. Lee, and J.-R. Zheng, "A Novel SIP-Based Route Optimization for Network Mobility," *IEEE J. Selected Areas Comm.*, vol. 24, no. 9, pp. 1682-1690, Sept. 2006.
- [20] M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "The Secure Real-Time Transport Protocol (SRTP)," IETF RFC 3711, Mar. 2004.
- [21] J. Arkko, E. Carrara, F. Lindholm, M. Naslund, and K. Norrman, "MIKEY: Multimedia Internet KEYing," IETF RFC 3830, Aug. 2004.
- [22] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, and J. Arkko, "Diameter Base Protocol," IETF RFC 3588, Sept. 2003.
- [23] S.-T. Wang, "SIP-Based Mobile VPN over Network Mobility (NEMO)," master's thesis, Nat'l Tsing Hua Univ., June 2007.
- [24] M. Garcia-Martin, M. Belinchon, M. Pallares-Lopez, C. Canales, and K. Tammi, "Diameter Session Initiation Protocol (SIP) Application," IETF RFC 4740, Nov. 2006.
- [25] P. Srisuresh, J. Kuthan, J. Rosenberg, A. Molitor, and A. Rayhan, "Middlebox Communication Architecture and Framework," IETF RFC 3303, Aug. 2002.
- [26] M. Handley and V. Jacobson, "SDP: Session Description Protocol," IETF RFC 2327, Apr. 1998.
- [27] J.-C. Chen and T. Zhang, *IP-Based Next-Generation Wireless Networks*. John Wiley and Sons, Jan. 2004.
- [28] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," IETF RFC 3550, July 2003.
- [29] J. Bilien, E. Eliasson, J. Orrblad, and J.-O. Vatn, "Secure VoIP: Call Establishment and Media Protection," *Proc. Second Workshop Securing Voice over IP*, June 2005.
- [30] D. Geneiatakis, T. Dagiuklas, G. Kambourakis, C. Lambrinouidakis, S. Gritzalis, S. Ehart, and D. Sisalem, "Survey of Security Vulnerabilities in Session Initiation Protocol," *IEEE Comm. Surveys Tutorials.*, vol. 8, no. 3, pp. 68-81, Apr.-June 2006.
- [31] D. Sisalem, J. Kuthan, and S. Ehlert, "Denial of Service Attacks Targeting a SIP VoIP Infrastructure: Attack Scenarios and Prevention Mechanisms," *IEEE Networks*, vol. 20, no. 5, pp. 26-31, 2006.
- [32] S. Salsano, L. Veltri, and D. Papalilo, "SIP Security Issues: The SIP Authentication Procedure and its Processing Load," *IEEE Networks*, vol. 16, no. 6, pp. 38-44, Nov. 2002.
- [33] D. Geneiatakis, G. Kambourakis, C. Lambrinouidakis, T. Dagiuklas, and S. Gritzalis, "A Framework for Protecting a SIP-Based Infrastructure against Malformed Message Attacks," *Computer Networks*, vol. 51, no. 10, pp. 2580-2593, July 2007.
- [34] D. Geneiatakis and C. Lambrinouidakis, "An Ontology Description for SIP Security Flaws," *Computer Comm.*, vol. 30, no. 6, pp. 1367-1374, Mar. 2007.
- [35] E. Rescorla, *SSL and TLS: Designing and Building Secure Systems*. Addison Wesley, 2001.
- [36] T. Taleb, H. Nishiyama, N. Kato, and Y. Nemoto, "Securing Hybrid Wired/Mobile IP Networks from TCP-Flooding Based Denial-of-Service Attacks," *Proc. IEEE GLOBECOM*, pp. 2907-2911, Dec. 2005.
- [37] D. Geneiatakis, G. Kambourakis, and T. Dagiuklas, "A Framework for Detecting Malformed Messages in SIP Networks," *Proc. 14th IEEE Workshop Local and Metropolitan Area Networks*, Sept. 2005.
- [38] J. Xie and I.F. Akyildiz, "A Novel Distributed Dynamic Location Management Scheme for Minimizing Signaling Costs in Mobile IP," *IEEE Trans. Mobile Computing*, vol. 1, no. 3, pp. 163-175, July-Sep. 2002.
- [39] W. Ma and Y. Fang, "Dynamic Hierarchical Mobility Management Strategy for Mobile IP Networks," *IEEE J. Selected Areas Comm.*, vol. 22, no. 4, pp. 664-676, May 2004.

- [40] R. Rummler, Y.W. Chung, and A.H. Aghvami, "Modeling and Analysis of an Efficient Multicast Mechanism for UMTS," *IEEE Trans. Vehicular Technology*, vol. 54, no. 1, pp. 350-365, Jan. 2005.
- [41] S. Fu, M. Atiqzaman, L. Ma, and Y.-J. Lee, "Signaling Cost and Performance of SIGMA: A Seamless Handover Scheme for Data Networks," *Wireless Communications and Mobile Computing*, vol. 5, no. 7, pp. 825-845, Nov. 2005.
- [42] Y.-B. Lin, "Reducing Location Update Cost," *IEEE/ACM Trans. Networks*, vol. 5, no. 1, pp. 25-33, Feb. 1997.
- [43] D. Gross and C.M. Harris, *Fundamentals of Queueing Theory*. John Wiley and Sons, 1998.
- [44] P. Calhoun, T. Johansson, C. Perkins, T. Hiller, and P. McCann, "Diameter Mobile IPv4 Application," RFC 4004, Aug. 2005.
- [45] "The Network Simulator - ns-2," <http://www.isi.edu/nsnam/ns>, 2011.
- [46] M.M. Zonoozi and P. Dassanayake, "User Mobility Modeling and Characterization of Mobility Patterns," *IEEE J. Selected Areas Comm.*, vol. 15, no. 7, pp. 1239-1252, Sept. 1997.
- [47] Y. Fang and I. Chlamtac, "Teletraffic Analysis and Mobility Modeling of PCS Networks," *IEEE Trans. Comm.*, vol. 47, no. 7, pp. 1062-1072, July 1999.



is a student member of the IEEE.



Tuan-Che Chen received the PhD degree from the Department of Computer Science, National Tsing Hua University, Taiwan, in 2010. He was a visiting student in the School of Electrical and Computer Engineering, Cornell University, New York, from 2009 to 2010. He was an intern at Telcordia Technologies, New Jersey, in 2006. His research interests include mobility management, energy efficiency of wireless networks, network coding, and performance evaluation. He

Jyh-Cheng Chen received the PhD degree from the State University of New York at Buffalo in 1998. He is currently a professor in the Department of Computer Science at National Chiao Tung University (NCTU), Hsinchu, Taiwan. He has been on the faculty at NCTU since August 2010. Dr. Chen was with Bellcore/Telcordia Technologies, Morristown, New Jersey, from 1998-2001 and Telcordia Technologies, Piscataway, New Jersey, from 2008-2010. He has also been with the Department of Computer Science, National Tsing Hua University (NTHU), Hsinchu, Taiwan, since 2001 as assistant/associate/full/adjunct professor. He is a coauthor of *IP-Based Next-Generation Wireless Networks* (Wiley, 2004), has published more than 80 papers, and holds 19 US, six ROC, and four PRC patents. He received the 2000 Telcordia CEO Award, the 2001 SAIC ESTC Publication Award, the 2004 NTHU New Faculty Research Award, the 2006 NTHU Outstanding Teaching Award, and the 2007 Best Paper Award for Young Scholars from the IEEE Communications Society Taipei and Tainan Chapters and the IEEE Information Theory Society Taipei Chapter. He is a technical editor of the *IEEE Wireless Communications* and was a guest editor of the *IEEE Journal on Selected Areas in Communications* special issue on all-IP wireless networks. He was the technical program cochair of MWCN 2007 and ITRE 2005 and was on the technical program committee of IEEE INFOCOM 2005-2006, IEEE GLOBECOM 2005-2011, and IEEE ICC 2007-2011. He was a tutorial speaker at IEEE GLOBECOM 2002, 2003, and 2006, and leads the development of WIRE1x. He is a senior member of the IEEE, the IEEE Computer Society, and the ACM.



Zong-Hua Liu received the PhD degree from the Department of Computer Science, National Tsing Hua University, Taiwan, in 2010. He was an intern at Telcordia Technologies, New Jersey, in 2009. His research interests include mobility management, admission control, resource management, and performance evaluation of wireless networks. He is a student member of the IEEE.