International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.9 (2014), pp.171-188
http://dx.doi.org/10.14257/ijmue.2014.9.9.19

Fast and Secure Session Mobility in IMS-based Vertical Handover Scenario

Bongkyo Moon

Department of Computer Science and Engineering, Dongguk Univ-Seoul, Korea
bkmoon@dongguk.edu

Abstract

In this paper, the make-before-break handover method is basically considered in order to reduce the SIP (session initiation protocol) session restoration delay due to the interface switching during WiFi-to-3G vertical handover under IMS (IP multimedia subsystem). Actually, the session restoration delay in WiFi-to-3G handover scenario can be dramatically reduced by performing IMS registration and session re-setup beforehand via WiFi link before actual vertical handover. However, the proper SA (security association) setup scheme between MH (mobile host) and P-CSCF (proxy call session control function) should be developed for the successful pre-authenticated registration since the IP address bound to 3G interface has to be used instead of the IP address bound to WiFi interface as the source IP address of the packets that the MH generates. In this paper, we have enhanced SIP header in order to solve this kind of IP address-mismatch problem. That is, a SA setup scheme is proposed for pre-authenticated registration in WiFi-to-3G handover scenario. We also define IMS handover delay as the sum of authenticated registration delay and session re-setup delay occurring under IMS-based vertical handover. We finally show that this delay reduced by the proposed scheme is acceptable enough to provide delay-sensitive real-time services. Moreover, the modification is deployed only at MH and P-CSCF, which is independent of other party's network as well as the CSCFs in MH's HN (home network).

Keywords: *IMS, vertical handover, security association, delay, enhanced SIP*

1. Introduction

IMS (IP multimedia subsystem) actually uses the underlying IP network as a universal communication infrastructure, and is hereby deployed in various environments such as stationary, mobile, wired and wireless, regardless of the type of access devices [1,4-5]. Currently, most smart phones (*e.g.*, iPhone and Android devices) are equipped with multiple wireless interfaces (*e.g.*, Bluetooth, 3G, Wi-Fi) and thus one smart phone can be associated with multiple IP addresses at any particular instance. Recent natural trend for smart phone users in vertical handover scenario between 3G and Wi-Fi is to utilize high-bandwidth in hotspots, and is to switch to 3G networks whenever user goes out of Wi-Fi coverage or link condition is not stable enough [3, 8].

Typically, the larger size of signaling message over wireless link with low bandwidth will decrease the link efficiency and thus degrades service quality. Unfortunately, the message size of text-based protocol, SIP (session initiation protocol), becomes larger than that of other binary protocols. Hence, SIP signaling efficiency in IMS has become increasingly important issue for providing interactive multimedia service such as real-time online gaming or VoIP service in 3G networks. However, it is a challenge to keep signaling delay low in IMS since SIP has text-based nature [4-5, 13].

ISSN: 1975-0080 IJMUE

Meanwhile, 3G-to-WiFi vertical handover actually experiences much less delay over wireless link since Wi-Fi link has much higher bandwidth than 3G and thus complex retransmission scheme such as Radio Link Protocol (RLP) is not required over Wi-Fi link. However, WiFi-to-3G handover may incur unacceptable delay for supporting real-time multimedia services since 3G radio access network is so vulnerable to noise as to increase the bit-error rate (BER) over the wireless channel. Thus, a semi-reliable link-layer retransmission mechanism such as RLP should be used to improve the BER performance over 3G wireless links [6, 14, 22]. This is particularly significant in the presence of lossy, time-variable and capacity-constrained wireless links.

Therefore, it is a major challenge to reduce the delay in transmitting SIP messages over the 3G wireless link for session re-setup at handover. Recently a proxy agent-based scheme is proposed to minimize the SIP session setup delay over a wireless link in 3G handover scenarios [13]. This scheme is based on the two characteristics. One is that the major factor of SIP session re-setup delay is generally caused by the retransmissions in the unreliable wireless links, and the other is that most of the fields in request messages as well as response messages are duplicated when a set of SIP messages are exchanged during session re-setup procedure. Moreover, IETF has developed a method for compressing SIP signaling message called SigComp [16]. TCCB (Text-based Compression using Cache and Blank approach) has been deployed for compressing SIP message between SIP clients and a proxy server in 3G network [15].

Nevertheless, it is still a critical point to reduce the SIP session restoration delay due to WiFi-to-3G vertical handover and keep it within a desirable maximum limit for interactive multimedia service under IMS (IP multimedia subsystem). Recently, link-layer assisted SIP mobility scheme is actually introduced for reducing WiFi-to-3G handover delay by sending SIP session re-setup message in advance via Wi-Fi link before actual vertical handover to 3G [2-3, 12]. However, this scheme does not consider vertical handover under IMS scenario. Hence, the make-before-break handover scheme needs to be considered under IMS, which performs registration and session re-setup via Wi-Fi link for the purpose of establishing 3G link connection in advance. For the proper SA (security association) setup between mobile host (MH) and the proxy server (P-CSCF) in the pre-authenticated registration procedure, however, the IP address bound to 3G interface should be used instead of the IP address bound to Wi-Fi interface for the source IP address of the packets MH sends to the P-CSCF.

In this paper, therefore, we have enhanced SIP header in order to handle this kind of IP address-mismatch problem for proper SA setup in IMS-based vertical handover scenario. That is, a SA setup scheme is proposed for pre-authenticated registration in WiFi-to-3G handover scenario. We also define IMS handover delay as the time for the authenticated registration and session re-setup procedures occurring at vertical handover under IMS. We show that IMS handover delay is reduced by the aid of the proposed scheme. In this scheme, no change is required in the SIP message processing except for the behaviors in both P-CSCF and MH. In section 2, IMS architecture and vertical handover scenario are investigated, and the enhanced SIP mechanism is introduced in section 3. Performance evaluations and discussions with results are given in section 4 and section 5, respectively. Conclusions are presented in section 6.

2. Authenticated Registration and Security Association in IMS

2.1. IMS Registration and Session Setup

IP Multimedia Subsystem (IMS), which architecture is shown in Figure 1, consists of various SIP servers called Call Session Control Functions (CSCFs), which performs the

multimedia session control, the address translation function, the voice coder negotiation for audio communications, and the management of the subscriber's profile [4-5]. More specifically, the proxy CSCF (P-CSCF) is the mobile's first point of contact in the IMS, the serving CSCF (S-CSCF) is responsible for the session management, and the interrogating CSCF (I-CSCF) is responsible for finding the appropriate S-CSCF based on load or capability. Hence, all SIP messages that MH transmits are first sent to P-CSCF, which then forwards them to another CSCF in MH's Home Network (HN) (*i.e.*, I-CSCF or S-CSCF). Similarly, all SIP messages transmitted toward MH are sent to P-CSCF, which then forwards them to MH. Also, S-CSCF maintains information such as MH's IP address and multimedia sessions.

Meanwhile, as shown in Figure 1, the user equipment like mobile host must share its physical address with the registrar in the network. That is, the user's public identity needs to be bound to the physical address along with registration step. Hence, the physical address can be changed with many times as a mobile host (MH) moves around the network, so the binding of public URI address may change frequently. The registration process actually begins when MH accesses IP network and obtains its IP address from the network. Once MH has obtained its IP address, SIP application is launched and MH then sends its address information to the SIP registrar [4-5].

In SIP session setup process, moreover, the calling party, user agent client (UAC), starts the transaction by sending a SIP INVITE request to the called party, user agent server (UAS). The INVITE request contains the details of the type of session that is requested and goes through the P/I/S-CSCF of the respective domains (see Fig. 1). Upon reception of the INVITE, the UAS sends its media parameters to the UAC. Then, the UAC decides on the proposed media parameters and returns its answer back to the UAS. When the called party decides to accept the call (*i.e.*, picks up), a 200 OK response is sent to the caller. The final step is to confirm the media session with an acknowledgment request ACK. Then, the media session is established [1, 4].

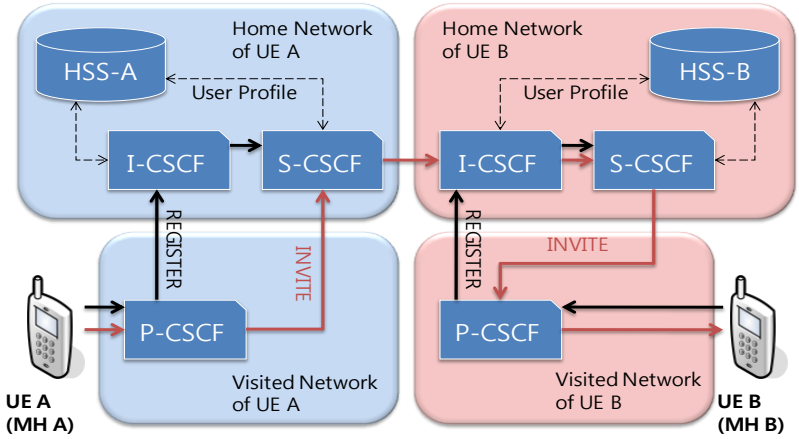

Figure 1. IMS Architecture and Registration

2.2. IMS-level Authenticated Registration

The IP Multimedia Subsystem (IMS) is essentially based on several security relations. Two of them, which are the authentication between user and network and the Security Association (SA) between the MH and the proxy Call Session Control Functions (P-CSCF), have an influence on SIP signaling. Actually, authentication and SA establishment procedures

in the IMS are directly coupled to SIP registration procedures. That is, SA is a relationship between two or more entities that describes how the entities will use security services to communicate securely. More specifically, IMS authentication is based on a shared secret and a sequence number (SQN), which is only available in the Home Subscriber Server (HSS) and IMS Service Identity Module (ISIM) application that is located in user's phone. Moreover, security via network interface is achieved by means of IPSec SAs, which require specific handling at the SIP signaling level. As the establishment of IPSec SAs is based on authentication of the user, new SAs are established during every re-authentication process. Consequently, new pairs of IPSec SAs have to be established between the MH and the P-CSCF.

IMS-level registration essentially includes the authentication procedure where IMS user requests authorization to use the IMS services in the IMS network. For IMS services, that is, IMS subscriber on MH should send SIP REGISTER request to the serving CSCF (S-CSCF) in its HN for registering its point of presence (*i.e.* MH's IP address). When MH initially attaches to new access network, mutual authentication between MH and HN needs to be accomplished within the initial IMS-level registration procedure. This kind of IMS-level registration triggering authentication is called authenticated registration [5, 9]. Unlike regular SIP procedure, registration within the IMS is mandatory before IMS subscriber can establish a session. Initial registrations are always authenticated, but other registrations may or may not be authenticated, depending on a number of security issues in the IMS. Only REGISTER request are authenticated, and other SIP requests, such as INVITE, are never authenticated by the IMS.

Meanwhile, the registration procedure in IMS actually completes after two round-trips, as illustrated in Figure 2. IMS users are authenticated by the S-CSCF with data provided by the HSS (home subscriber server), which is not shown in detail in Figure 2. That is, the S-CSCF receives the REGISTER request and authenticates the user. Specifically, the S-CSCF creates a SIP '401 Unauthorized' response, which includes a challenge that the MH should answer and then forwards it to the MH, via the I-CSCF and P-CSCF. The MH produces an appropriate response (known as credentials) to that challenge, and then sends a new SIP REGISTER request to the P-CSCF, which actually does the same operation as for the first REGISTER request. Lastly, the S-CSCF sends a 200 (OK) response to the REGISTER request, to indicate the success of the REGISTER request after validating user credentials via the HSS.

The detail SIP message (SM) flow for the authenticated registration and security association (SA) setup procedure is also depicted in Fig. 2. In this figure, regarding SM1 REGISTER request and SM6 '401 Unauthorized' response, MH becomes UAC and P-CSCF works as UAS. Meanwhile, P-CSCF becomes UAC and I-CSCF works as UAS with regard to SM2 REGISTER request and SM5 '401 Unauthorized' response, and so on [5]. When SM1 is transmitted, MH adds 'Via' header into SM1. The 'Via' header indicates where to send the response for the request, and thus P-CSCF shall send SM6 response to the address of 'Via' header contained in SM1 (*i.e.*, MH's IP address) [5].

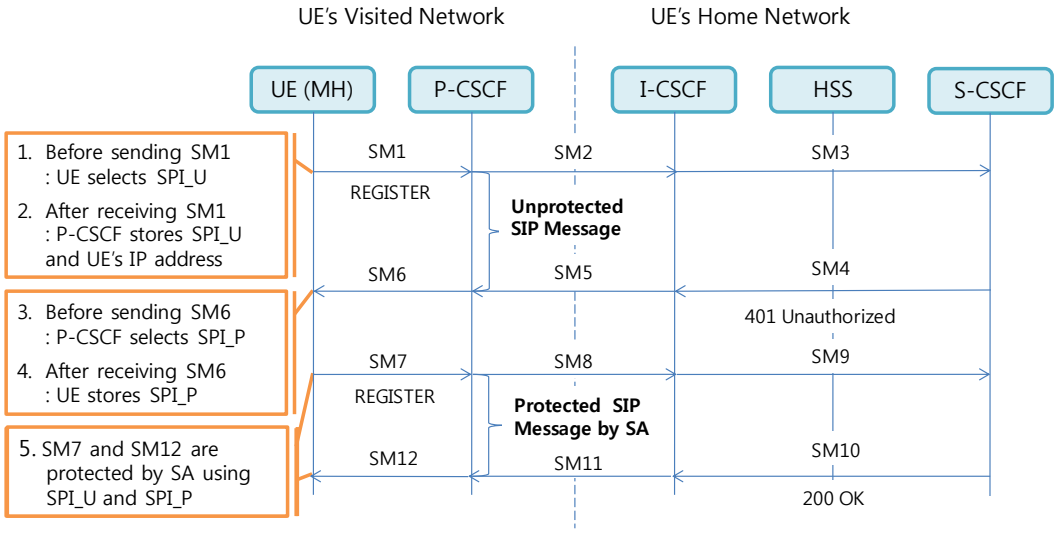

Figure 2. IMS Registration Procedure with Security Association (SA) Setup

2.3. Security Association Setup between MH and P-CSCF

During authenticated registration in IMS, SIP messages exchanged between the MH and the P-CSCF should be protected by Security Association (SA). There should be at least two security connections for both directions since the SA is unidirectional for each pair of communicating systems. The SA is uniquely identified by a randomly chosen unique number called the security parameter index (SPI) and the IP address of the destination. That is, when a system sends a packet that requires IPSec protection, it looks up the SA in its database, applies the specified processing, and then inserts the SPI from the SA into the IPSec header. When the IPSec peer receives the packet, it looks up the SA in its database by the destination address and SPI and then processes the packet as required. SA is simply a statement of the negotiated security policy between the MH and the P-CSCF.

2.3.1 SA Setup within Registration Procedure: SPI is locally allocated for SAs. In an authenticated registration, the MH and the P-CSCF each select two SPIs for the new SAs, which have not been associated with existing SAs. Moreover, IP addresses are bound to two pairs of SAs for inbound SA and outbound SA at the P-CSCF. That is, in the case of inbound SA at the P-CSCF, the source and destination IP addresses associated with the SA are identical to those in the header of the IP packet received by the P-CSCF. In the case of outbound SA at the P-CSCF, the source IP address equals the destination IP address bound to the inbound SA, and the destination IP address equals the source IP address bound to the inbound SA [5, 17].

During the authenticated registration procedure in Fig. 2, SPI_U and SPI_P are actually selected by MH and P-CSCF, respectively. And the P-CSCF and the MH agree on a set of parameters to establish the two IPSec SAs between them. When MH sends SM1 (REGISTER request) containing the selected SPI_U to P-CSCF, P-CSCF stores SPI_U and MH's IP address indicated in SM1 header. Later, P-CSCF then sends SM6 containing the selected SPI_P to MH. Upon receipt of SM6, MH can establish SA with P-CSCF. In this step, the P-CSCF obtains the integrity and encryption keys in a '401 Unauthorized' response sent from the S-CSCF, and then removes both keys from the response before relaying it to the MH. The

P-CSCF and the MH use the same two REGISTER transactions that are used for authentication to negotiate the rest of the IPSec parameters.

Consequently, when MH registers new IP address, which is bound to its network interface, at S-CSCF in its HN, IMS registration procedure can be completed only if SA setup succeeds between the MH and the P-CSCF. That is, after new SA is well established, MH is successfully registered with the new IP address. Once SA is established, all SIP messages exchanged between MH and P-CSCF are securely protected by the SA.

2.3.2 SA Establishment and Protected Ports

Without establishing SAs between MH and P-CSCF, the P-CSCF is typically allowed to receive only REGISTER messages and error messages on unprotected ports, all other messages arriving on the unprotected port shall be either discarded or rejected by the P-CSCF. Similarly, since the MH is allowed to receive only the responses to unprotected REGISTER messages and error messages on an unprotected port, all other messages arriving on a unprotected port shall be rejected or silently discarded by the MH. Eventually, one SA is established from the MH's client-protected port to the P-CSCF's server-protected port, and the other SA goes from the P-CSCF's client-protected port to the MH's server-protected port. Both SAs support traffic in both directions [9, 17].

More specifically, the set of new SAs actually needs to be established with a shared key. That is, when the MH and the P-CSCF establish two IPSec SAs between them, they need to agree on shared keys, which are obtained from the IPSec SA parameters, for protecting SIP signaling between them (RFC3329). Unfortunately, the P-CSCF knows nothing about the security parameters that are shared between user's ISIM application and the HSS in the home network. In IMS registration step, however, the S-CSCF actually sends the integrity key (IK) and the ciphering key (CK) to the P-CSCF in the 401 (Unauthorized) response. The P-CSCF must remove these two keys from the header and store them locally before sending the 401 (Unauthorized) response toward the MH. The IK is then used by the P-CSCF as the shared key for the set of SAs. The MH on the other side of the network interface calculates the IK from the received challenge in the 401 (Unauthorized) response and also uses it as the shared key. By means of the IK, the P-CSCF and the MH can then establish the set of SAs between the four ports beforehand in the initial REGISTER request and its response [18].

Figure 3 shows the protected ports and the SAs with TCP between MH and P-CSCF. In this figure, one SA is established from the MH's client-protected port to the P-CSCF's server-protected port and the other goes from the P-CSCF's client-protected port to the MH's server-protected port. Both SAs support traffic in both directions. That is, MH and P-CSCF using TCP between them send responses on the same TCP connection (*i.e.*, using the same ports) as they received the request.

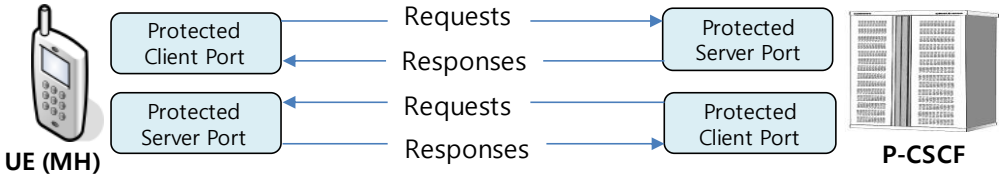

Figure 3. Ports and Security Associations with TCP between MH and P-CSCF

The P-CSCF and the MH use the same two REGISTER transactions (shown in Figure 2) that are used for authentication to negotiate the rest of the IPSec parameters. The following headers show the examples of the fields added in REGISTER and the RESPONSE messages

by MH and P-CSCF, respectively. That is, *Security-Client* header field that the MH adds to the REGISTER (SM1) contains the mechanisms (*ipsec-3gpp*) and algorithms (hmac-sha-1-96) that the MH supports as well as the SPIs and port numbers that it uses.

Security-Client : ipsec-3gpp; alg=hmac-sha-1-96;
spi-c=23456789; spi-s=12345678;
port-c=2468; port-s=1357

Also, *Security-Server* header field that the P-CSCF adds to the 401 (Unauthorized) response (SM6) contains the mechanisms (*ipsec-3gpp*) and algorithms (hamc-sha-1-96) that the P-CSCF supports as well as the SPIs and port numbers that it uses. The SAs are ready to be used as soon as the MH receives the *Security-Server* header field (SM6). So, the MH sends a REGISTER (SM7) request over one of the just established SAs.

Security-Server : ipsec-3gpp; q=0.1; alg=hmac-sha-1-96;
spi-c=98765432; spi-s=87654321;
port-c=8642; port-s=7531

3. SA Setup Scheme for Pre-Authenticated Registration

3.1. Vertical Handover Scenario under IMS

Typically, when MH with two physical interfaces performs vertical handover between 3G and Wi-Fi, SIP user agent (UA) on MH can be associated with two IP addresses at any particular instance. During vertical handover event, however, the UA finally disconnects an existing SIP session bound on one interface after completing SIP session re-setup via the other interface. This decision depends upon the destination route metrics or local policy. In case of 3G-to-WiFi vertical handover scenario, on-going connection can be kept constantly through either 3G or Wi-Fi since 3G network usually covers Wi-Fi hot spots. In WiFi-to-3G scenario, however, it is not easy to estimate exactly when user goes out of Wi-Fi coverage (Figure 3). Hence, the existing service might be disrupted unless new 3G link connection can be completely established before MH's Wi-Fi connection becomes not valid any more.

In order to keep on-going SIP session constantly after WiFi-to-3G vertical handover event, more specifically, the IP address associated with the picked-up 3G interface eventually needs to be registered and maintained within the S-CSCF in HN. And then MH should send SIP re-INVITE message containing new IP address bound to 3G interface to each one of its corresponding hosts (CHs). That is, MH re-invites the CHs to its new temporary address by sending INVITE message through several P/I/S-CSCF servers. The re-INVITE message actually uses the same call identifier as in the original call setup, and also includes MH's original SIP user identifier and its new IP address for the purpose of informing the CH where MH wants to receive future SIP messages [1, 4]. Once CH gets the updated information about the MH, it sends an acknowledge message while starting to send data to the MH. In a typical mid-session vertical handover scenario, consequently, total IMS handover delay is mainly caused by the SIP message exchanged for the SIP location update and SIP re-INVITE request after MH attaches to 3G access network.

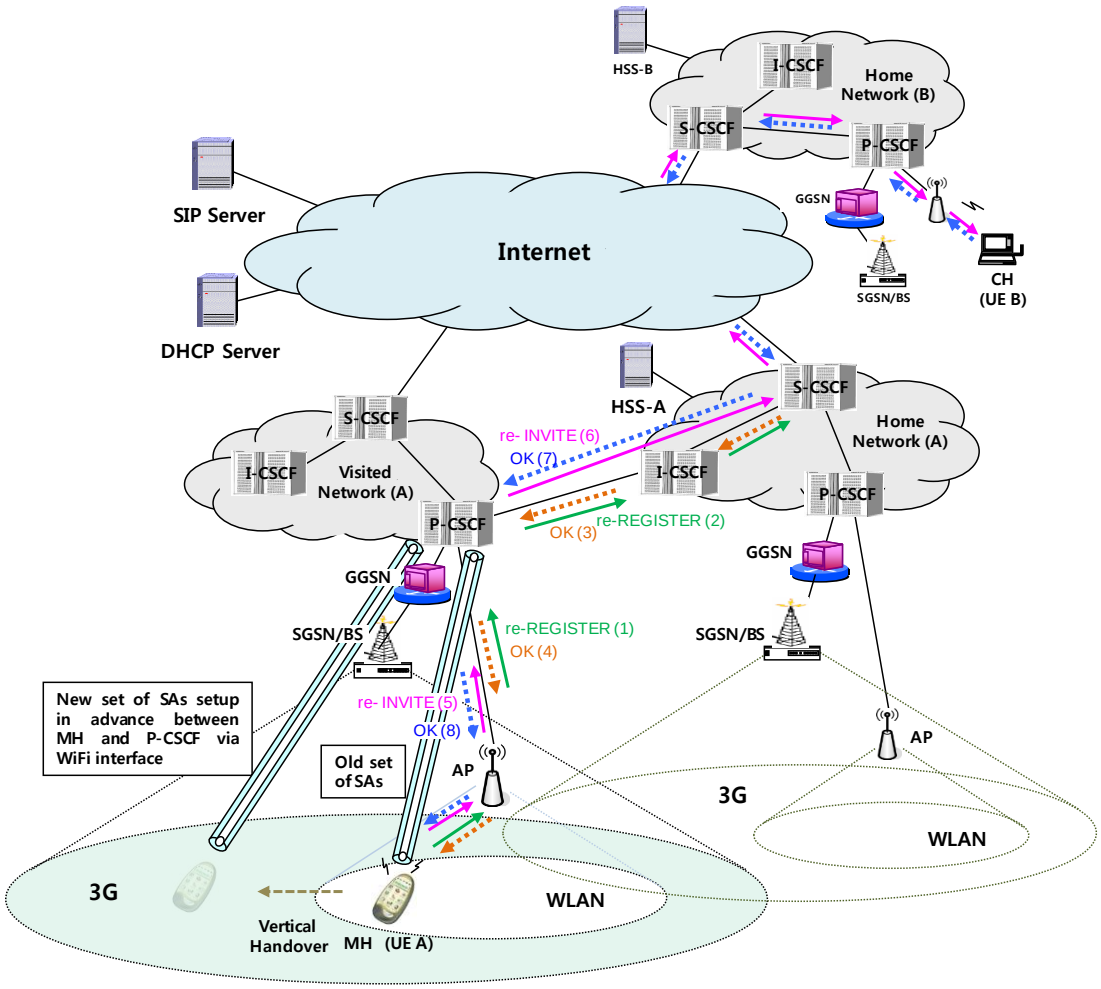

Figure 4. Pre-Authenticated Registration and Session re-Setup Procedures for WiFi-to-3G Vertical Handover

3.2. The Proposed Scheme

Several mechanisms [2-3, 12] have actually been introduced for reducing WiFi-to-3G handover delay by performing registration and session re-setup with new IP address in advance via Wi-Fi link before actual vertical handover to 3G, where new IP address is actually bound to MH's 3G interface. That is, since REGISTER message is sent in advance via Wi-Fi link instead of 3G interface to the S-CSCF in MH's home network, vertical handover delay can be much reduced.

Figure 4 shows the WiFi-to-3G vertical handover scenario in tightly-coupled 3G and Wi-Fi interworking architecture under IMS. It is assumed that MH is equipped with both Wi-Fi and 3G interfaces and stays in visited network (VN) as IMS subscriber. It is also assumed that MH has been initially communicating with CH attached to the network as mobile node. We here define IMS handover delay as the time taken for the authenticated registration and session re-setup at handover event in IMS. It is typically necessary for new mechanism to be developed for reducing IMS handover delay at WiFi-to-3G vertical handover. In this section, hence, a SA setup scheme is proposed for pre-authenticated registration and session re-setup in WiFi-to-3G handover scenario.

Figure 4 also illustrates the SIP message exchanges for re-establishing the connection between the MH and the CH by performing pre-registration for fast WiFi-to-3G vertical handover while the MH is in the visited network. For this kind of mechanism under IMS, REGISTER and re-INVITE messages are actually exchanged via Wi-Fi interface between the MH and the P-CSCF in its visited network.

We here define IMS handover delay as the time taken for the authenticated registration and session re-setup at handover event in IMS. It is typically necessary for new mechanism to be developed for reducing IMS handover delay at WiFi-to-3G vertical handover. In this section, hence, a SA setup scheme is proposed for pre-authenticated registration and session re-setup in WiFi-to-3G handover scenario.

Table 1. Examples of (a) traditional SIP messages exchanged via 3G link after actual vertical handover and (b) enhanced SIP messages exchanged via Wi-Fi link before actual vertical handover (In this table, it is assumed that MH's IP addresses for 3G and Wi-Fi are 1.2.3.4 and 5.6.7.8, respectively, and the P-CSCF's IP addresses for 3G and Wi-Fi are a.b.c.d and e.f.g.h, respectively).

```
REGISTER sip:proxy.wonderland.com SIP/2.0
Via: SIP/2.0/UDP 1.2.3.4:5060;branch=z9hG4bKjjf9d45
Max-Forwards: 70
To: <sip:alice@wonderland.com>
From: <sip:alice@wonderland.com>;tag=635529
Call-ID: 99183245223553@43je8ew9236
CSeq: 540 REGISTER
Contact: <sip:alice@1.2.3.4>
Expires: 6300
Content-Length: 0
```

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP 1.2.3.4:5060;branch=z9hG4bKjjf9d45
To:<sip:alice@wonderland.com>;tag546229
From: <sip:alice@wonderland.com>;tag=635529
Call-ID: 99183245223553@43je8ew9236
CSeq: 540 REGISTER
Contact: <sip:alice@1.2.3.4>
```

a) Traditional SIP messages via 3G link after actual vertical handover

```
REGISTER sip:proxy.wonderland.com SIP/2.0
Via: SIP/2.0/UDP 5.6.7.8:5060;branch=z9hG4bKjjf9d45
Max-Forwards: 70
To: <sip:alice@wonderland.com>
From: <sip:alice@wonderland.com>;tag=635529
Call-ID: 99183245223553@43je8ew9236
CSeq: 540 REGISTER
Contact: <sip:alice@1.2.3.4>
Src-for-SA: 1.2.3.4
Expires: 6300
Content-Length: 0
```

```
SIP/2.0 200 OK
Via: SIP/2.0/UDP 5.6.7.8:5060;branch=z9hG4bKjjf9d45
To:<sip:alice@wonderland.com>;tag546229
From: <sip:alice@wonderland.com>;tag=635529
Call-ID: 99183245223553@43je8ew9236
CSeq: 540 REGISTER
Contact: <sip:alice@1.2.3.4>
Dst-for-SA: a.b.c.d
```

b) Enhanced SIP messages via WiFi link before actual vertical handover

3.2.1. New Header Fields in SIP Message

For the proper SA setup between MH and the P-CSCF in the pre-authenticated registration step, the IP address bound to 3G interface should be used instead of the IP address bound to Wi-Fi interface for the source IP address of the packets that MH sends to the P-CSCF. Therefore, we have enhanced SIP header for handling this kind of IP address mismatch problem for SA setup in IMS-based vertical handover scenario. That is, a SA setup scheme is here proposed using enhanced SIP headers.

We actually define two new header fields (*i.e.*, Src-for-SA and Dst-for-SA) in SIP message and modify UAC behavior at MH and UAS behavior at P-CSCF. The additional behaviors of the modified UAC are to generate SIP request containing Src-for-SA header and to process SIP response containing Dst-for-SA header. On the other hand, the behaviors of the modified UAS are to process SIP request containing Src-for-SA header and to generate SIP response containing Dst-for-SA header. Specifically, Src-for-SA header field is included in SIP request message sent from MH to P-CSCF. Thus, P-CSCF uses Src-for-SA for SA setup via 3G link instead of the source IP address in the packet header of the request message

transmitted via Wi-Fi link. Similarly, SIP response message includes Dst-for-SA header field. That is, Dst-for-SA is used for SA setup via 3G link instead of the destination IP address in the packet header of the response message transmitted via Wi-Fi link from P-CSCF to MH. Table 1 presents the examples of enhanced SIP messages exchanged via Wi-Fi link before actual vertical handover and traditional SIP messages exchanged via 3G link after actual vertical handover.

3.2.2. Modified UAC Behavior at MH

- Generating SIP request. Before sending SIP request, MH checks Wi-Fi link status using the information obtained from layer 2. Actually the interaction with layer 2 is out of focus in this paper and handover scheme using link layer information was studied in [2]. If Wi-Fi connection is still valid when the request message is generated, the enhanced SIP would be performed on MH. That is, three steps are actually taken on MH. First, IP address for 3G interface is inserted into the 'Contact' header. Second, Src-for-SA header that contains IP address for 3G interface is added. Lastly, this request is transmitted toward P-CSCF via Wi-Fi (Algorithm1: line2~line5). If Wi-Fi connection is not valid when the request message is generated, this request is sent toward P-CSCF via 3G according to traditional SIP.

- Processing SIP response. If Dst-for-SA header is included in the received message, Dst-for-SA header shall be used for SA (Algorithm 1: line10~line11). Otherwise, destination IP address in packet header of the response shall be used for SA setup.

Algorithm 1 Modified UAC Behavior at MH

```
1:  if  SIP request message is generated then
2:      { set "Contact: MH's IP for 3G"}
3:      if WiFi connection is still valid then
4:          {add "Src-for-SA: MH's IP for 3G"}
5:          {transmit request message via WiFi}
6:      else
7:          {transmit request message via 3G}
8:      end;
9:  else if SIP response message is processed then
10:     if Dst-for-SA header is included in the message then
11:         {use Dst-for-SA for SA setup}
12:     else
13:         {use destination IP address in packet header for SA}
14:     end;
15: end;
```

3.2.3 Modified UAS Behavior at P-CSCF

- Processing SIP request. If the received message includes Src-for-SA header, P-CSCF performs the enhanced SIP (Algorithm 2: line2~ line5). Since the value of 'Via' header (*i.e.*, MH's IP address for 3G interface) is different from the source IP address (*i.e.*, MH's IP address for WiFi interface) in the transmitted packet header, P-CSCF adds 'received' field into the 'Via' header. The value of 'received' field is actually source IP address of the

transmitted packet header. Since the 'Via' header in enhanced SIP doesn't indicate the actual response path, consequently, P-CSCF shall send response message to IP address in the 'received' field [4]. Next, P-CSCF shall use Src-for-SA header for SA setup. Lastly, P-CSCF stores the information about the request message, such as Call-ID and sequence number, for sending the response message back. This procedure for enhanced SIP is setup with 'e-SIP flag'.

- **Generating SIP response.** If the e-SIP flag is already enabled when SIP response message is generated, P-CSCF shall perform the enhanced SIP (Algorithm 1: line10~line12). That is, P-CSCF adds Dst-for-SA header with MH's IP address for 3G interface and then transmits this response message to MH via Wi-Fi. Otherwise, the normal SIP response message is transmitted to MH via 3G.

Algorithm 2 Modified UAS Behavior at P-CSCF

```
1:  if SIP request message is processed then
2:      if Src-for-SA header is included in the message then
3:          {set "received: MH's IP for WiFi"}
4:          {use Src-for-SA for SA setup}
5:          {set "e-SIP Flag"}
6:      else
7:          {use source IP address in packet header for SA}
8:      end;
9:  else if  SIP response message is generated then
10:     if "e-SIP Flag" is enabled then
11:         {add "Dst-for-SA: P-CSCF's IP for 3G"}
12:         {transmit response message to MH via WiFi}
13:     else
14:         {transmit response message to MH via 3G}
15:     end;
16:  end;
```

4. Performance Evaluations

In this section, we present the delay analysis model for the IMS session setup signaling procedures with an emphasis on the IMS registration process. We assume that both UEs are mobile in our scenario but only one UE takes part in vertical handover event at the moment. For the simplicity, abbreviated in the Figure 2 is the procedure which I-CSCF sends a Diameter User-Authentication-Request (UAR) to the HSS (home subscriber server) for authorization and determination of S-CSCF already allocated to the user. Internet delay is not considered here since it can be regarded as a constant.

In order to evaluate the SA setup scheme proposed for pre-authenticated registration in WiFi-to-3G handover scenario, we define the IMS handover delay, D_{IMS} which consists of registration delay, $D_{REGISTER}$ and session re-setup delay, D_{INVITE} in IMS. $D_{REGISTER}$ takes two RTTs (round trip time) between MH (UE_A) and the S-CSCF in MH's HN. On the other hand, D_{INVITE} is the one-way delay since CH (UE_B) can send data to MH's new address as soon as it receives re-INVITE message. To compute D_{IMS},

we consider queuing delay and transmission delay. In the following subsection, these delays are described in detail.

$$D_{IMS} = D_{REGISTER} + D_{INVITE} \tag{1}$$

$$\begin{aligned} D_{REGISTER} = 2 \cdot \big(& D_{UE\_A} + D_{wireless\_A} + D_{P-CSCF} \\ & + D_{I-CSCF} + D_{S-CSCF} + D_{I-CSCF} \\ & + D_{P-CSCF} + D_{wireless\_A} + D_{UE\_A} \big) \end{aligned} \tag{2}$$

$$\begin{aligned} D_{INVITE} = & D_{UE\_A} + D_{wireless\_A} + D_{P-CSCF} \\ & + D_{S-CSCF} + D_{I-CSCF} + D_{S-CSCF} \\ & + D_{P-CSCF} + D_{wireless\_B} + D_{UE\_B} \end{aligned} \tag{3}$$

4.1. Queueing Delay

The SIP session setup may take considerable time due to SIP message processing delay in MH, intermediate servers (*e.g.*, P-CSCF, I-CSCF, S-CSCF) and destination server. The major delay can be roughly estimated using the queueing theory based on waiting time formulas. Actually, the MH and the CSCF servers perform dedicated jobs, but the destination server may serve a variety of non-SIP related tasks as well as the SIP messages. In order to compute the queueing delay, therefore, an M/M/1 queueing model can be deployed for the MH and the CSCF servers, and a priority based M/G/1 model can be derived for the destination server. Hence, according to M/M/1 queueing model for MH (UE_A) and CSCFs and M/G/1 model for CH (UE_B), the queuing delay estimates at MH, CSCFs and CH are given as follows:

$$D_{UE-A} = \frac{1}{\mu - \lambda_A} \tag{4}$$

$$D_{P-CSCF} = D_{I-CSCF} = D_{S-CSCF} = \frac{\rho_s}{\lambda(1-\rho_s)} \tag{5}$$

$$D_{UE-B} = \frac{\frac{1}{\mu_s}(1-\rho_o-\rho_s)+R}{(1-\rho_o)+(1-\rho_o-\rho_s)} \tag{6}$$

SIP message arrival rate at MH, λ_A is a fraction of SIP message arrival rate at CSCF servers, λ: $\lambda_A \leq \lambda$. Thus, the average queueing delay at MH is given as the formula (4), where μ is the service rate of SIP message at MH. The average queueing dealys at the P/I/S-CSCF follow the identical formula (5), where ρ_s is destination and CSCF server's loads. The queueing delay at the destination is the formula (6), where ρ_o is the load at the destination for non-SIP messages and μ_s is the service rate of SIP messages at the destination. The value R equals to $\frac{\lambda_o \bar{x}_o^2 + \lambda_s \bar{x}_s^2}{2}$ where $\bar{x}_o^2$ and $\bar{x}_s^2$ are the second moments of μ_o and μ_s, respectively.

4.2. Transmission Delay

When messages are transmitted over 3G link, Radio Link Protocol (RLP) is usually used in order to overcome erroneous wireless link. Hereby, 3G wireless link introduces major delays in comparison with the queueing and transmission delay over the backbone networks. That is,

IMS delay for SIP session re-setup at WLAN-to-3G vertical handover is mainly limited when SIP signaling messages are transmitted over erroneous and bandwidth-limited 3G wireless link. In order to compute the delay occurred when SIP messages are transmitted over wireless link, hence, the delay model for frame and packet transmission can be used [6, 14].

The assumptions and analytic method in this section heavily rely on the previous works [6-7]. For the analysis of transmission delay with RLP, several parameters need to be defined. p is the probability of an RLP frame being in error in the air link. C_{ij} represents the first frame received correctly to the destination at the *i*th retransmission of the *j*th retransmission trials. That is, the missing frame has been lost up to the (*j*-1)th retransmission trial and up to the (*i*-1)th retransmissions in the *j*th trial. Hence, the probability of transmitting a frame successfully at the *i*th retransmission of the *j*th retransmission trials after frame transmission error is given as

$$\mathrm{P}(\mathrm{C}_{\mathrm{ij}}) = \mathrm{p}(1-\mathrm{p})^2((2-\mathrm{p})\mathrm{p})^{\frac{\mathrm{j}(\mathrm{j}-1)}{2}+\mathrm{i}-1} \quad . \tag{7}$$

Therefore, the probability of transmitting a frame successfully over the air link with RLP operating underneath is given as

$$\mathrm{P}_{\mathrm{f}} = 1-\mathrm{p}+\sum_{\mathrm{j}=1}^{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{j}}\mathrm{P}(\mathrm{C}_{\mathrm{ij}}) = 1-\mathrm{p}(\mathrm{p}(2-\mathrm{p}))^{\frac{\mathrm{n}(\mathrm{n}+1)}{2}} \quad , \tag{8}$$

where n is the maximum number of RLP retransmission trials. Hence, packet loss rate or the probability of retransmission (*q*) is computed as follows:

$$q = 1 - P_f^{\,k} \quad , \tag{9}$$

where the parameter *k* is the number of frames in a packet transmitted over the air. Considering the RLP retransmissions, the delay in transmitting a packet containing k frames over the RLP is given by

$$D' = D + (k-1)\tau + \frac{k\left(P_f - (1-p)\right)}{P_f^2}\left(\sum_{j=1}^{n}\sum_{i=1}^{j}P(C_{ij})\left((2j)D + \left(\frac{j(j+1)}{2}+i\right)\tau\right)\right), \tag{10}$$

where D is the end-to-end frame propagation delay over the air link and τ is the interfame time of RLP frame. Since the SIP messages are assumed to be sent over TCP, D_{3G}, the delay to transmit a TCP segment consisting of k frames over 3G wireless link with RLP is given by

$$D_{3G} = D' + \frac{2Dq(1-q)}{1-q^{Nm}}\left(1+\frac{4q\left(1-(2q)^{Nm-2}\right)}{1-2q}-\frac{q\left(1-q^{Nm-2}\right)}{1-q}\right), \tag{11}$$

where N_m is the number of TCP retransmissions.

Since Wi-Fi link with high bandwidth does not require retransmission protocol such as RLP, meanwhile, transmission delay without RLP, D_{WiFi} is given as follow:

$$D_{WiFi} = (k-1)\tau + \frac{D}{(1-q^{Nm})(1-2q)} + \frac{D(1-q)}{1-q^{Nm}}\left(\frac{q^{Nm}}{1-q}-\frac{2^{Nm+1}q^{Nm}}{1-2q}\right), \tag{12}$$

where $q = 1-p^k$, *k* is the number of frames in a packet transmitted over the Wi-Fi, τ is the interfame time, and D is the end-to-end frame propagation delay over the Wi-Fi.

5. Evaluation Results and Discussions

In this section, we present the numerical results for the delay analysis of SIP session re-setup signaling under IMS-based vertical handover scenario.

5.1. SA Setup Scheme via 3G Link after Actual Vertical Handover

We evaluate IMS handover delay under various network conditions (i.e., FER and SIP session arrival rate) in WiFi-to-3G vertical handover scenario. After acquiring a new IP address bound to 3G interface, MH sends SIP messages via 3G interface. That is, $D_{wireless\_A}$ becomes D_{3G}. Under the assumption that SIP message size is 500 bytes, hence, the value of k for D_{3G} in each channel can be derived as k=21 for 9.6 kbps, k =11 for 19.2 kbps and k =2 for 128 kbps. In order to focus on vertical handover at MH (UE_A) side, we assume that CH (UE_B) is just connected to Wi-Fi. That is, $D_{wireless\_B}=D_{WiFi}$ is assumed. In this analysis, the SIP message arrival rate at CSCF (λ) is given as $10\lambda_A$ and the service rate of SIP message at MH (μ) is 4×10^{-4}. The service rate of SIP message at CH (μ_s) is also given as μ and the load for non-SIP messages at CH (ρ_o) is 0.7. The server load on CSCF and CH (ρ_s) is given as λ/μ and R is $0.501[\rho_o^2+\rho_s^2]$. End-to-end frame propagation delay (D) is given as 100ms and the interframe time (τ) is 20ms. The number of TCP retransmissions (N_m) is given as 10 and the maximum number of RLP retransmission trials (n) is also given as 3.

Figure 5 shows IMS handover delay with normal SIP as FER is increased, where λ_A=50 requests/s. Figure 6 presents IMS handover delay with varying λ_A at MH side where FER is 0.05. In Figure 5 and Figure 6, the minimum IMS handover delay over 128 kbps is nearly kept as 0.6s (=600ms). Despite that L2 and L3 handover latencies are not considered here, this IMS handover delay is unacceptable for real-time applications, which require handover delay less than 100ms and at most not more than 200ms.

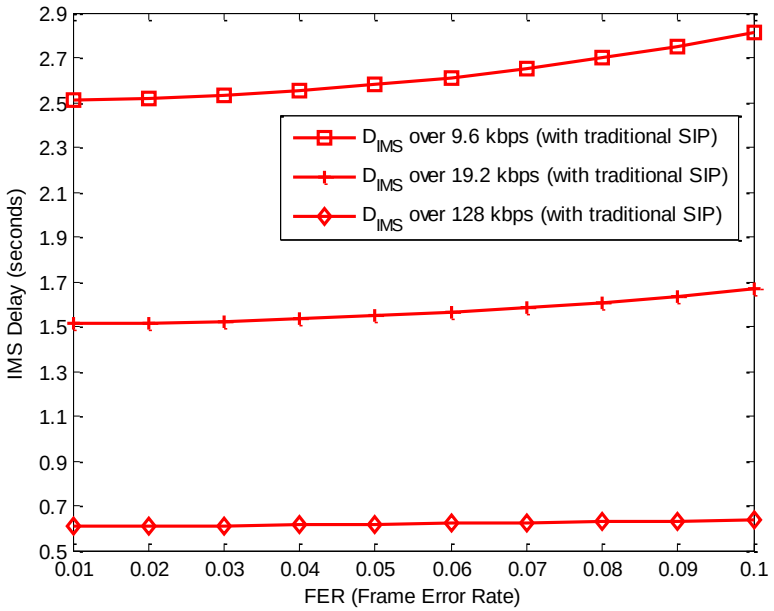

Figure 5. IMS Handover Delay vs. FER in SA Setup Scheme via 3G Link

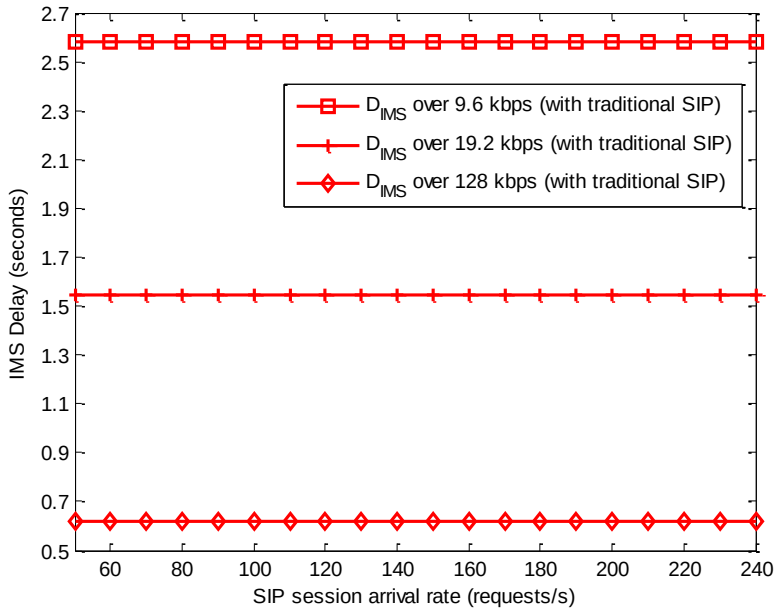

Figure 6. IMS Handover Delay vs. Session Arrival Rate in SA Setup Scheme via 3G Link

5.2. SA Setup Scheme via Wi-Fi Link before Actual Vertical Handover

In WiFi-to-3G handover scenario, the enhanced SIP messages for authenticated registration and session re-setup are sent via Wi-Fi interface instead of 3G. That is, $D_{wireless\_A}$ becomes D_{WiFi}. Also, due to the additional headers, the size of the enhance SIP message becomes a little larger than normal SIP message. Since Wi-Fi channel has relatively high bandwidth, however, this additional header size is trivial and can be neglected. That is, k=1 for 2 and 11Mbps. In this analysis, τ and N_m are given as 1ms and 10, respectively. D is also given as 0.27ms and 0.049ms for 2Mbps and 11Mbps Wi-Fi channel, respectively.

Figure 7 and Figure 8 show the IMS handover delay with the enhanced SIP under same network conditions as Figure 5 and Figure 6, respectively. In Figure 7 and Figure 8, the maximum IMS handover delay via 2Mbps channel is nearly 11.6ms. On the other hand, the maximum IMS handover delay via 11Mbps Wi-Fi is nearly 10ms. In comparison with the results in Figure 5 and Figure 6, consequently, we can know that the IMS handover delay, if the enhanced SIP is deployed, is acceptable enough to provide real-time applications for the user on mobile device experiencing WiFi-to-3G handover.

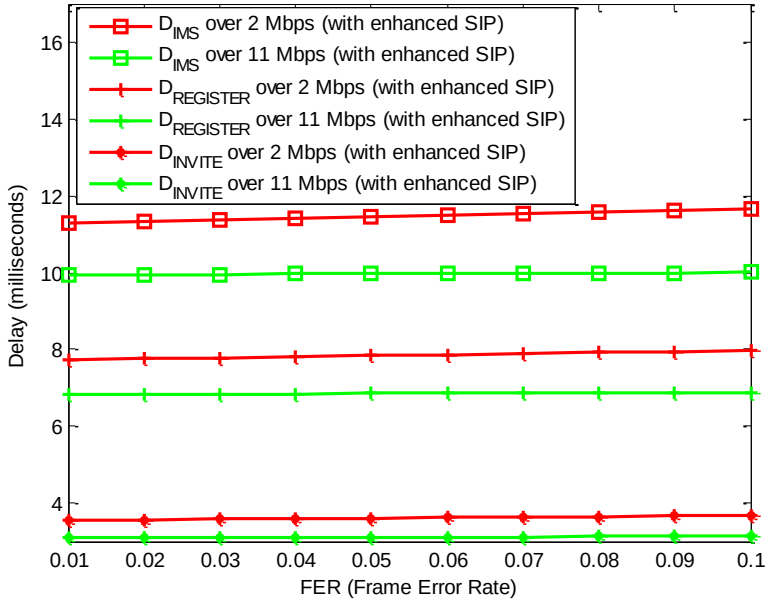

Figure 7. IMS handover delay vs. FER in SA setup scheme via WiFi link

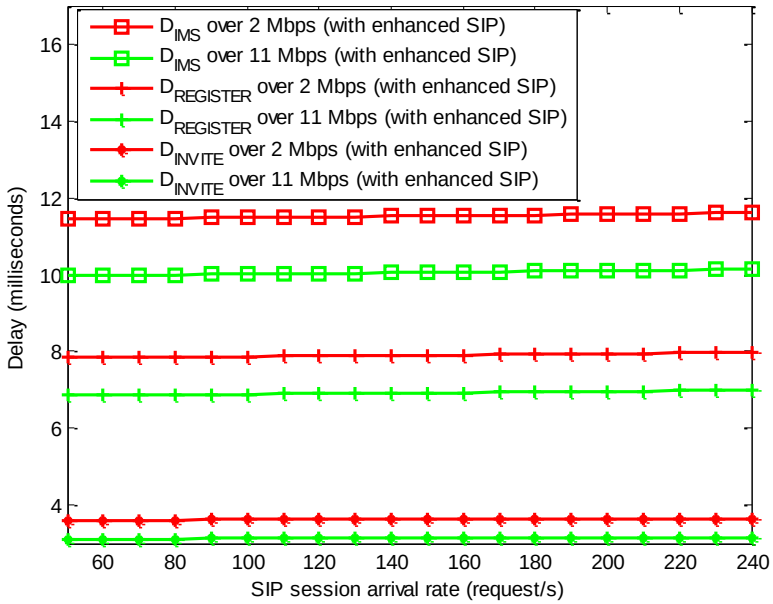

Figure 8. IMS Handover Delay vs. Session Arrival Rate in SA Setup Scheme via Wi-Fi Link

6. Conclusion

In this paper, we investigated that the IMS delay at WiFi-to-3G vertical handover may cause considerable service disruption. However, this delay can be dramatically reduced by performing IMS registration and session re-setup beforehand via Wi-Fi link

before actual vertical handover. Actually, it is essential to setup SAs between MH and P-CSCF in the authenticated registration step under IMS-based scenario. Consequently, a SA setup scheme with enhanced SIP was proposed in order to solve the IP address-mismatch problem caused by the WiFi-to-3G interface switching for the successful pre-authenticated registration. The results showed that IMS handover delay reduced by the proposed scheme is acceptable enough to support delay-sensitive real-time applications. Moreover, the modification is deployed only at MH and P-CSCF, which is independent of other party's network as well as CSCFs in MH's HN.

Acknowledgement

This work was supported by the research program of Dongguk University. Many special thanks are also given to Hoyeon Lee for her help on this research work.

References

[1] J. Rosenberg and H. Schulzrinne, "SIP: Session Initiation Protocol," IETF RFC3262, **(2002).**
[2] W. Kim, M. Kim, K. Lee, C. Yu and B. Lee, "Link Layer Assisted Mobility Support Using SIP for Real-time Multimedia Communications," MobiWac'04 **(2004)**, pp.127-129.
[3] J. McNair and F. Zhu, "Vertical Handoffs in Fourth-Generation Multinetwork Environments," IEEE Wireless Communications, pp.8-15, **(2004).**
[4] H. Sinnreich and A. Johnston, "Internet Communications Using SIP – Delivering VoIP and Multimedia Services with Session Initiation Protocol," 2nd Edition. Wiley Publishing, Inc. **(2006).**
[5] G. Camarillo and M. A. Garcia-Martin, "The 3G IP Multimedia Subsystem (IMS)," 2nd edition. Wiley **(2006).**
[6] H. Fathi, S. Chakraborty and R. Prasad, "Optimization of SIP Session Setup Delay for VoIP in 3G Wireless Networks," IEEE Transactions on Mobile Computing, vol. 5, no. 9, **(2006),** pp. 1121-1132.
[7] H. Fathi, S. Chakraborty and R. Prasad, "On SIP Session Setup Delay for VoIP Services Over Correlated Fading Channels," IEEE Transactions on Vehicular Technology, vol. 55, no.1, **(2006)** January.
[8] T. Ahmed, K. Kyamakya and M. Ludwig, "Architecture of a Context-Aware Vertical Handover Decision Model and Its Performance Analysis for GPRS-WiFi Handover," The 11th IEEE Symposium on Computers and Communications (ISCC'06) **(2006)**, pp. 795-801.
[9] "Access security for IP-based services (Release 8)," 3GPP TS33.203 v8.1.0 **(2007).**
[10] M. Melnyk, A. Jukan, and C. Polychronopoulos, "A Cross-Layer Analysis of Session Setup Delay in IP Multimedia Subsystem (IMS) With EV-DO Wireless Transmission," IEEE Transactions on Multimedia, vol. 9, no. 4, **(2007).**
[11] A. Munir and A. Gordon-Ross, "SIP-Based IMS Signaling Analysis for WiMax-3G Interworking Architectures," IEEE Trans. on Mobile Computing, vol. 9, no. 5, **(2010),** pp. 733-750.
[12] X. Yan, Y. A. Şekercioğlu and S. A Narayanan, "Survey of vertical handover decision algorithms in Fourth Generation heterogeneous wireless networks." Computer Networks, vol. 54, no. 11, **(2010),** pp. 1848-1863.
[13] B. Moon, "Minimizing SIP Session Re-Setup Delay over Wireless Link in 3G Handover Scenarios," EURASIP Journal on Wireless Communications and Networking, doi:10.1155/2010/634810 **(2010).**
[14] N. Banerjee, W. Wu, K. Basu and S. K. Das, "Analysis of SIP-based Mobility Management in 4G Wireless Networks," Computer Communications, vol. 27, **(2004),** pp. 697-707.
[15] I. Majumdar, V. Kenneally and D. Pesch, 'Improving SIP Call Control Performance through Message Compression - The TCCB Algorithm," SIP 2003 **(2003)**, Paris, France,
[16] H. Wook and S. Kang, "Improvement of link efficiency by compressing SIP signaling messages with SigComp," ICACT 2008 **(2008)**, pp. 1314-1317.
[17] Z. Wang, "IMS Security Framework," 3GPP2 S.S0086-B, Version: 2.0 **(2008).**
[18] M. Poikselka, G. Mayer, H. Khartabil and A. Niemi, "The IMS: IP Multimedia Concepts and Services in the Mobile Domain," Wiley **(2004).**
[19] W. Stallings, "Network Security Essentials," 3rd Ed., Pearson **(2007).**
[20] O. Hersent, "IP Telephony – Deploying VoIP Protocols and IMS Infrastructure," Wiley **(2011).**
[21] M. Poikselka, H. Holma, J. Hongistro, J. Kallio and A. Toskala, "Voice over LTE (VoLTE)," Wiley **(2012).**
[22] B. Moon, "Analysis of Ongoing SIP Session with Resource Reservation in Vertical Handover Scenario," Wireless Personal Communications, doi:10.1007/s11277-013-1160-6 **(2013).**

Author

Bongkyo Moon, he received the B.S. degree in Computer Science from Sogang University, Korea, in 1992, the M.S. degree in information and communications from GIST (Gwangju Institute of Science and Technology), Korea, in 1998, and the Ph.D. degree in Telecommunications from KCL (King's College London), London, UK. He worked as a researcher in software and telecommunication areas at INEX Technologies, Inc., Santa Clara, CA, USA from 1992 to 1996 and at ETRI (Electronics and Telecommunications Research Institute), Korea, from 1998 to 1999. He also worked a senior researcher in the Telecommunication R and D Centre, Samsung Electronics, Korea, from 2003 to 2005. Since 2005, he has been working as faculty member (currently associate professor) in dept. of Computer Science and Engineering, Dongguk University, Seoul, Korea. His research interests are mobile computing, cloud computing and network security.