Seamless Session Mobility Scheme in Heterogeneous Wireless Networks

Kuan-Lin Chiu¹, Yuh-Shyan Chen*†, and Ren-Hung Hwang¹

¹ Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi 621, Taiwan, R.O.C.
² Department of Computer Science and Information Engineering, National Taipei University, San-Shia, Taipei 237, Taiwan, R.O.C.

SUMMARY

The next generation wireless communication system will likely be heterogeneous networks, as various technologies can be integrated on heterogeneous networks. A mobile multiple-mode device can easily access the Internet through different wireless interfaces. The mobile multiple-mode device thus could switch to different access points to maintain the robustness of the connection when it can acquire more resources from other heterogeneous wireless networks. The mobile multiple-mode device therefore needs to face the handover problem in such environment. This work introduces Session Initiation Protocol (SIP)-based cross-layer scheme to support seamless handover scheme over heterogeneous networks. The proposed scheme consists of a battery lifetime based handover policy and cross-layer fast handover scheme, called the SIP-based mobile stream control transmission protocol (SmSCTP). This work describes the major idea of the proposed scheme and infrastructure. The proposed scheme has been implemented in Linux system. The simulation and numerical results demonstrate that the proposed SmSCTP scheme yields better signaling cost, hand-off delay time, packet loss and delay jitter than SIP and mSCTP protocols. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: Handover, cross-layer design, session mobility, heterogeneous wireless networks

1. Introduction


*Correspondence to: Yuh-Shyan Chen, Department of Computer Science and Information Engineering, National Taipei University, San-Shia, Taipei 237, Taiwan, R.O.C.
†E-mail:yschen@mail.ntpu.edu.tw

Contract/grant sponsor: National Science Council of the Republic of China; contract/grant number: NSC-97-2219-E-305-001 and NSC-98-2219-E-305-001

Received 1 April 2010
Copyright © 2000 John Wiley & Sons, Ltd.
System (UMTS) [8] and High Speed Downlink Packet Access (HSDPA) [31], have been considered to provide ubiquitous communication services for a 4G wireless communication services. However, an acceptable solution for heterogeneous networks is the UMTS/HSDPA [9][10][11] network combined with the IEEE 802.11 WLAN [12] network; that is called WLAN/3.5G heterogeneous networks. In such heterogeneous networks, a mobile multiple-mode device can move between wireless networks. To maintain connectivity and robustness of the connection, the mobile multiple-mode device requires a fast and seamless mobility management scheme in end node. Therefore, how to develop a fast handover scheme with handover policy over heterogeneous networks is still a open issue.

In 3rd Generation Partnership Project (3GPP) specification document [5], SIP has been selected as the main signaling protocol in the UMTS/HSDPA IMS, as it can support multimedia service management and session mobility scheme [13]. A main component in the IMS, called the Call Session Control Function (CSCF), is defined to control an SIP session. However, SIP does not provide a seamless handover scheme to the mobile multiple-mode device. On the other hand, mobile SCTP (mSCTP) [14][34] is designed to support flexible association for wide application services in transport-layer using multihoming scheme. The multihoming scheme supports soft and seamless handover scheme. Therefore, combining SIP with mSCTP could be a good solution for WLAN/3.5G heterogeneous networks.

This work attempts to provide a session mobility scheme to mobile multiple-mode devices in heterogeneous networks by taking advantage of cross-layer design and can be compatible with current UMTS/HSDPA IMS framework. We concentrate on cross-layer design that uses the SIP protocol to perform handover decision and maintain handover procedure in application-layer, and utilizes the mSCTP scheme to switch different heterogeneous networks in transport-layer. To achieve seamless handover between the proposed WLAN/3.5G heterogeneous networks, we introduced a SIP-based mSCTP scheme, called as SmSCTP, which is a node-initiation scheme, in [32]. SmSCTP develops and specifics two new SIP signaling messages to realize the soft handover in proposed infrastructure. In addition, the proposed scheme investigates a handover policy to trigger SmSCTP. The Goal of the handover policy is to maximize goodput during a limited battery lifetime. Therefore, the mobile multiple-mode device can switch arbitrarily by using SmSCTP and acquire the maximum goodput before the battery is unavailable by the proposed handover policy.

SmSCTP is evaluated through NS-2 simulator. SmSCTP is also implemented in Kphone [30] software which works on Linux system. The simulation results show that SmSCTP achieves a fraction of handover latency, delay jitter and packet lost significantly. SmSCTP also provides better signaling cost than current SIP and mSCTP schemes.

The remainder of this paper is organized as follows. Section 2 discusses related work. Section 3 then presents the system model and underlying ideas. Next, Section 4 describes the handover policy on the proposed SmSCTP scheme and the details. Section 5 summarizes the performance analysis results. Conclusions are finally drawn in Section 6, along with recommendations for future research.

2. Related Work

Currently, handover schemes can be classified as soft or hard handover. The hard handoff [15][16][17] indicates that the ongoing connection is broken before a new connection is
established. During a hard handoff, a mobile multiple-mode device only communicates with one wireless network at a time. Conversely, the soft handoff \cite{18,19,20,21,22} refers to a mobile multiple-mode device connecting to multiple access networks simultaneously or when more than one connection exists between mobile multiple-mode devices.

Wu et al. \cite{17} presented an SIP-based hard handoff procedure called re-INVITE that provides session mobility on heterogeneous networks. Although this scheme supports seamless mobility on heterogeneous networks, handoff latency remains high. Sharma et al. \cite{16} investigated the mobile IP mechanism in network layers. This scheme must modify significantly the underlying networking infrastructure as well as mobile devices. Smadi et al. \cite{20} recently developed the Dynamically Anchored Conferencing Handoff (DACH) solution for soft handoff between WLAN and UMTS networks. The DACH algorithm uses the concept of an anchor to handle sessions. Azhari et al. \cite{18} proposed the Vertical Handoff Support Node (VHSN) that forwards and redirects a connection to a mobile device via a cellular network. When an abrupt connection is lost on a WLAN, the VHSN intercepts and redirects the media flow through the local cellular base station using its cellular end station attachment. Inoue et al. \cite{19} developed the Multimedia Integrated network by Radio Access innovation (MIRAI) server, which works as both a Basic Access Signaling (BAS) agent and mobile IP agent. However, the MIRAI server still suffers high handoff latency and high traffic overhead. Salsano et al. \cite{21} introduced the Back-to-Back User Agent (B2BUA), a bottleneck component, for managing sessions. Oh et al. \cite{22} presented a seamless handover scheme in IPv6 network.

In heterogeneous networks, the handover scheme also can be categorized as both vertical and homogeneous when switching to a different or the same wireless network. The environment in this work is based on the WLAN/UMTS heterogeneous networks. Therefore, homogeneous handover does not need to be discussed.

The vertical handoff \cite{23,24} is a popular approach for session mobility. The vertical handoff coordinates different wireless interfaces in a mobile device and can maintain a connection between two or more wireless heterogeneous networks. Snoeren et al. \cite{25} examined TCP-Migrate to allow an active connection across IP addresses. The TCP-Migrate mechanism requires changes to the TCP protocol for all hosts. Sur et al. \cite{26} proposed a framework based on cross layer rules in the Wideband Code Division Multiple Access (WCDMA) and WLAN wireless networks. They introduced a core component, called the Rule Engine (RE), to collect the information from multiple layers of a protocol stack or other databases. The RE uses these inputs and pre-defined rules to predict a handoff. Hence, the RE can provide a reliable vertical handoff; however, they are still hard handovers.

Recently, Kim et al. \cite{35} proposed a novel retransmission scheme for mSCTP to solve the packet reordering problem in heterogeneous wireless networks. While switching different SCTP connections, a mobile multiple-mode device can fast retransmit non-acknowledged packets through another high speed wireless interface, e.g. WLAN.

3. System Model and Basic Idea

This section describes the system model of this work. This model is based on 3GPP R6 documents and supports IMS services. The basic principle underlying the proposed SIP-based mSCTP scheme is then described in Section 3.2.
3.1. System Model

In heterogeneous networks, the UMTS/HSDPA network supports large transmission coverage to mobile multiple-mode devices. But, the WLAN network provides high transmission rate. The mobile multiple-mode device thus needs to switch different heterogeneous networks to acquire a good communication. Figure 1 shows the proposed architecture. In this architecture, many mobile multiple-mode devices connects to the UMTS/HSDPA and WLAN networks, individually. The mobile multiple-mode device and User Equipment (UE) (Fig. 1) are interchangeable in this work. Notably, a UE can move arbitrarily between the WLAN and UMTS/HSDPA networks. Each part is shown separately as follows.

In the 3GPP R99 specification, the UMTS network can be divided into two parts-circuit-switched (CS) and packet-switched (PS) service domains. The UE connects to Public Switched Telephone Network (PSTN) through the CS service. The PS service allows the UE to access the Internet services. The PS domain connects to an external Packet Data Network (PDN) through a Serving GPRS Support Node (SGSN) and a Gateway GPRS Support Node (GGSN) (Fig. 1(a)). The current position information of a user is stored in the SGSN, such that an incoming data packet is routed to the user immediately. The GGSN has a similar role as a gateway between the PDN and UMTS core network. Consequently, the GGSN carries out tasks associated with network access control, packet routing and mobility management. In terms of mobility management, the GGSN tunnels to the SGSN using the GPRS Tunnel Protocol (GTP). When a user moves into another Radio Network System (RNS) (Fig. 1(b)), which is controlled by the SGSN, the GGSN updates the SGSNID address and tunnels packets to the SGSN via the GTP.

The CSCF supports session control within the IMS (Fig. 1(c)). The CSCF server is an SIP server that is responsible for interacting with Home Subscriber Server (HSS) for mobility and session control. The CSCF server can be divided into three principal components-Proxy-CSCF...
SEAMLESS SESSION MOBILITY SCHEME IN HETEROGENEOUS WIRELESS NETWORKS

(P-CSCF), Serving-CSCF (S-CSCF) and Interrogating-CSCF (I-CSCF). The P-CSCF of the IMS is the first contact point for the UE. The main function of the P-CSCF is to forward SIP messages to the S-CSCF or I-CSCF depending on the message type or procedure. The S-CSCF is an important element for IMS. As the SIP proxy server, the S-CSCF forwards SIP messages to an appropriate node. The S-CSCF also supports the registration procedure. The I-CSCF functions as the contact point of an operator network. Similar to a doorkeeper, the I-CSCF supports charging, routing, and forwarding SIP messages to an appropriate node. Notably, the I-CSCF of this work is ignored for illustration ease (Fig. 1(c)). Figure 1(d) displays the interworking between the UMTS/HSDPA and WLAN networks. The wireless access gateway (WAG) [2] forwards data and messages from or to the PDN. In addition, the WAG can connect to more than one wireless access point.

According to characteristics of the UE and proposed architecture, this work makes the following assumptions.

1. A UE can always associate with the UMTS/HSDPA network prior to a WLAN network.
2. The MAC-layer of the UE provides physical and statistic information to upper layer.
3. The UE is authenticated by the UMTS/HSDPA authentication mechanism [27].

3.2. Basic Idea

This work highlights some of the major contributions to have handover policy and fast handover scheme in a UE. The handover policy is an application-layer algorithm, and the UE gathers and calculates the information of signal-to-noise ratio (SNR), transmission rate, and transmission probability from both wireless networks periodically. The handover policy also collects the information of the battery. Therefore, the UE can connect to a better access point to acquire the maximum transmission goodput depending on the battery lifetime. On the other hand, this work uses SmSCTP to achieve fast handover in proposed infrastructure. SmSCTP is a cross-layer scheme, which combines the SIP scheme at layer 7 and the mSCTP scheme at layer 4. In layer 7, to utilize the flexibility of the SIP, the SmSCTP scheme develops two signaling messages to increase and switch access points. In addition, the SmSCTP scheme provides a multi-homing mechanism using mSCTP in layer 4. Therefore, SmSCTP performs the soft handover procedure when SIP sends a handover request to mSCTP. Notably, the UE does not transmit control messages of the current mSCTP scheme to the corresponding node (CN). The control messages are replaced by the developed SIP messages.

Two novel SIP signaling messages are designed for controlling session mobility. One is Annexing IP, called as ANNEXIP, and the other is Switching Session, termed as SSESSION. The ANNEXIP message inserts a location information of a UE into the databases of the CSCF server and the CN simultaneously. Notably, the ANNEXIP message can be sent continuously once the UE acquires a new IP address from different wireless networks. Each new IP address can establish a new association between the UE and CN. Conversely, the SSESSION message confirms the final location of the UE. Therefore, upon receiving the SSESSION message, the CSCF server and CN set the primary location as recorded in the SSESSION message, and the CN then forwards the packets to the primary location. The CN also redirects data to this location. The SSESSION message is unlike the ANNEXIP message in that it is transmitted only once. According to the cross-layer design, the SmSCTP scheme cooperates with the mSCTP to perform add IP address (ADD-IP) and the handoff procedures. In addition, the SSESSION message is able to update the location information of the UE in the CSCF server.
and CN; thus, the SIP REGISTER message is cancelled.

4. Cross-layer Fast Handover Scheme over Heterogeneous Networks

This section illustrates more details of the proposed scheme. The handover policy is described in section 4.1. Then, the transaction and implementation of SmSCTP are illustrated in section 4.2 and 4.3, respectively.

4.1. Handover Policy

Before switching to other access points, a UE needs to perform handover policy algorithm as shown in Fig. 2. The goal of this algorithm is to maximize the transmission goodput during a limited battery lifetime. The algorithm consists of "full" and "emergent" phases. In full phase, the remainder battery is more than a threshold value, e.g. $RB \geq T_e$. The UE selects a better access point which supports high transmission rate in spite of the UE needs more transmission power to send packets. However, the UMTS/HSDPA and WLAN networks are the bandwidth sharing system. The UEs thus have to contend the wireless resources. Let $tp_i$ be the transmission probability, where $i$ be the wireless network type, which is calculated periodically by the UE. Then, the UE computes the maximum transmission rate of each wireless network, denoted as $C_i$. Once $C_{\text{wifi}} \geq C_{\text{hsdpa}}$, the UE handovers to the WLAN network. Otherwise, the UE switches to UMTS/HSDPA network.

In emergent phase, the battery lifetime of a UE is ended soon. Depending on the transmission rates assigned from different wireless networks, the UE has to acquire a maximum goodput before the battery is unavailable. The UE thus computes the maximum goodput, denoted as $g_i$, where $i$ be the wireless network type, during the remainder transmission time. Therefore, if $g_{\text{wlan}} \geq g_{\text{hsdpa}}$, the UE switches to WLAN network. The algorithm works continuously until the battery is unavailable. The analysis of aforementioned phases is shown as follows.

![Figure 2. Handover Policy for SmSCTP](#)
In full phase, for heterogeneous networks, it is difficult to use SNR value intuitively, because the transmission coverage of the UMTS/HSDPA network is greater than the WLAN network, but the transmission data rate of the WLAN network is higher than the UMTS/HSDPA network. Therefore, this work uses the available bandwidth, which driven by the Shannon capability theorem [33], to be a basis of the handover. The Shannon’s theorem with transmission probability for the WLAN network is stated as:

\[ C_i = b_i \cdot \log_2 \left( 1 + \frac{S_i}{N_i + 1} \right) \cdot t_p, \]  

where \( t_p = \frac{tt_{UE}}{f_A} \), subject to \( 0 \leq t_p \leq 1 \), \( tt_{UE} \) be the transmission time requested by a UE in a frame, \( t_A \) is a transmission frame time in an access point. Thus, \( t_p \) is a probability that the packet can be transmitted successfully by the UE. Additionally, \( b \) is the assigned bandwidth by the system, \( S \) is the signal power, and \( N \) is the noise.

In emergent phase, the remainder battery lifetime for the WLAN network (\( RBL_i \)) can be calculated by:

\[ RBL_i = \frac{RP}{p_i}, \]  

where \( RP \) be the remainder battery energy, and \( p_i \) be the transmission power per unit. Finally, the maximum goodput can be defined as follows:

\[ g_i = C_i \cdot RBL_i, \]  

Notable that the value of \( g_i \) is not fixed because the UE moves. Hence, the UE still gathers SNR and recomputes the value of \( g_i \).

4.2. Operation of SmSCTP

By applying new methods to the SIP protocol, the SmSCTP provides multi-homing capability to support smooth handovers. This section first describes the transaction of the SIP-based SmSCTP scheme. Although the SmSCTP scheme can be operated on a WLAN \( \rightarrow \) UMTS and UMTS \( \rightarrow \) WLAN simultaneously, where symbol \( \rightarrow \) denotes the moving direction. However, this section only addresses the operation of WLAN \( \rightarrow \) UMTS. Thus, the UE first connects to a WLAN, and then uses SmSCTP to switch the access network to the UMTS network.

At initiation, a UE is able to acquire new IP address from the system. Let \( IP_{CN} \) be an IP address of the CN, where \( IP_{CN} \in Pool_{CN} \). Let \( IP_i \) be an IP address, where \( i \geq 1 \) and \( IP_i \in Pool_{UE} \), which is assigned from the Dynamic Host Configuration Protocol (DHCP) server. Notably, the SmSCTP scheme enables to set up more than one connection once the UE associates with different wireless networks. Each connection can be defined as \( (IP_{CN}, IP_i) \).

The UE selects a best connection with the batter goodput as the primary. On the other hand, the UE executes the handover policy periodically before handover. In each phase of handover policy, the UE can switch access point when the SSESSION procedure is triggered.

Figure 3 shows the operations of the SmSCTP, which consists of three stages. Figure 3(a) shows the Packet Data Protocol (PDP) context procedure of the UMTS network. In this stage, a UE is communicating with the CN through the WLAN, and an SIP session has been established. Hence, the primary connection is \( <UE \leftrightarrow AP \leftrightarrow WAG \leftrightarrow PDN \leftrightarrow CN> \).
During the call, the UE collects signal strength from the UMTS/HSDPA network. When associating with the UMTS/HSDPA network successful, the UE performs the PDP context activation procedure to attach an immediate SGSN. The PDP context activation request is then sent to a GGSN to acquire an IP address and network resources. Upon receiving this IP address, the UE adds the assigned IP address to $Pool_{UE}$. The details of this procedure are shown as follows.

**Step 1:** The UE sends an Activate PDP Context Request message to an immediate SGSN when the UE has detected and associated with a UMTS/HSDPA RAN. The SGSN then negotiates with an immediate RNS to allocate resources for the UE.

**Step 2:** The immediate SGSN sends a Create PDP Context Request message to a GGSN.
to establish a GTP tunnel between the immediate SGSN and GGSN. At this time, the GGSN allocates an IP address and resources for the UE.

**Step 3:** The GGSN replies with a Create PDP Context response message to the immediate SGSN when the GTP tunnel is established. The assigned UMTS/HSDPA IP address and information regarding resources for the UE are included in this message.

**Step 4:** Finally, The SGSN returns an Activate PDP Context Accept message, which includes the UMTS IP address and resource information, to the UE. The UE adds the assigned IP address to $Pool_{UE}$.

Figure 3(b) shows the ANNEXIP procedure in the proposed scheme. After the PDP context procedure is complete, the UE obtains a new IP address from the UMTS/HSDPA network. The UE then utilizes the ANNEXIP message to notify the CN to configure a new connection using the new IP address. In this procedure, the primary connection remains $<UE \leftrightarrow AP \leftrightarrow WAG \leftrightarrow PDN \leftrightarrow CN>$, and the backup connection is $<UE \leftrightarrow RNS \leftrightarrow SGSN \leftrightarrow GGSN \leftrightarrow PDG \leftrightarrow PDN \leftrightarrow CN>$.

**Step 5:** The UE sends a ANNEXIP message, which includes a new UMTS/HSDPA IP address information, to the CSCF server. The CSCF server initiates a timer for this ANNEXIP message to maintain the session state.

**Step 6:** The CSCF server inserts a via header in top of the SSESSION message to ensure that the reply message can be through by the CSCF server. The CSCF server then forwards the ANNEXIP message to the CN. Upon receiving message, the CN adds IP information to the peer IP pool. Hence, two connections have been established between the UE and CN; the primary connection is $<UE \leftrightarrow AP \leftrightarrow WAG \leftrightarrow PDN \leftrightarrow CN>$, and the backup connection is $<UE \leftrightarrow RNS \leftrightarrow SGSN \leftrightarrow GGSN \leftrightarrow PDG \leftrightarrow PDN \leftrightarrow CN>$.

**Step 7:** The CN replies the 200 OK message sent to the CSCF server when the CN establishes the association. The CSCF server then cancels the timer.

**Step 8:** The CSCF server forwards the 200 OK message to UE to complete the ANNEXIP procedure.

Figure 3(c) shows the SSESSION procedure. In this procedure, the UE continuously performs the handover policy. When the handover policy sends a handover request to SmSCTP, the UE sends the SSESSION message to the CN to exchange the main and backup paths. After receiving the SSESSION message, the CSCF server updates the location information of the UE and the CN re-directs data packets to new primary connection. Therefore, the primary connection is changed from $<UE \leftrightarrow AP \leftrightarrow WAG \leftrightarrow PDN \leftrightarrow CN>$ to $<UE \leftrightarrow RNS \leftrightarrow SGSN \leftrightarrow GGSN \leftrightarrow PDG \leftrightarrow PDN \leftrightarrow CN>$.

**Step 9:** The UE sends the SSESSION message to the CSCF server through the UMTS/HSDPA core network. The SSESSION message includes an IP address, which will be the primary. The CSCF server updates the location information of the UE and sets another timer to maintain session status.
Step 10: The CSCF server inserts a via header in top of the SSESSION message. The CSCF server then forwards the SSESSION message to the CN. Upon receiving this message, the CN resets the primary connection and demands mSCTP to redirect the session from previous primary connection to new primary connection on-the-fly.

Step 11: Completed aforementioned steps, the CN replies to the 200 OK message to the UE through the CSCF server. The CSCF server then expires the timer. The primary connection in this step becomes \(<\text{UE} \leftrightarrow \text{RNS} \leftrightarrow \text{SGSN} \leftrightarrow \text{GGSN} \leftrightarrow \text{PDG} \leftrightarrow \text{PDN} \leftrightarrow \text{CN}>\).

4.3. Implementation of the SmSCTP

Figures 4(a) and (b) show the ANNEXIP and SSESSION messages, respectively. This work implemented the SmSCTP protocol in Linux. According to the RFC-3261 and 3GPP TS23.228 documents, the new signaling messages do not modify the original SIP message structure. The black dashed box in Figs. 4(a) and (b) denote modified and inserted headers. The Request-Line shows the goal of such messages. This work uses the Request-Line header to display the name of the new SIP method. In session description protocol (SDP), this work inserts a new header to a newly acquired IP address or final IP address. In the ANNEXIP message, if a UE can obtain more than one IP address, each new IP address results a new ANNEXIP message is sent. On the other hand, the IP address of the exchanged primary path is recorded in the SDP of the SSESSION message. The header names in the ANNEXIP and SSESSION messages are the same as the message name.

![Figure 4. Header content of new SIP messages. (a) ANNEXIP (b) SSESSION](image)
5. Performance Analysis
This section assesses the performance of the proposed SmSCTP mathematically and using the developed testbed. The effectiveness of the proposed approach is compared with that of two schemes—the SIP re-INVITE scheme developed by [13] and the mSCTP scheme developed by [36]. The performance metrics used were signaling cost, handoff latency, packet loss, and delay jitter. The signaling cost and handover latency were determined mathematically in Sections 5.1 and 5.2, respectively. On the other hand, this work represents the mobility pattern of a UE by using call mobility ratio (CMR). Additionally, the signaling cost and handover latency are also compared with implementation results. In Sections 5.3 and 5.4, an NS-2 simulator is used determine the packet loss and delay jitter.

Figure 5 depicts the implementation environment for SmSCTP. The SmSCTP, SIP and mSCTP schemes are implemented on the testbed. In the 3G UMTS system, Taiwan’s Chunghwa Telecom (CHT) system is used. The CHT system provides download links at 384 Kbps and uplinks at 64 Kbps. The CHT operator can assign a public IP address to the UE component. The IEEE 802.11b network is also used to support WLAN connectivity. In the UE, Kphone [30] software running on Linux Fedora system. The Kphone is an SIP soft phone based on the RFC-3261 standard, and this work modified this software to support the functionalities of the proposed scheme. In addition, this work also modifies the layer-4 SCTP socket and to replace the UDP socket. The four performance metrics are defined as follows.

- **Signaling cost** measures the number of signaling message during a call. The control signaling is counted when the UE handed over.
- **Handover latency** is the time taken by the handoff procedure. This time does not include transmitting time of the end-to-end data packet.
- **Packet loss** is defined as the number of packets not receive through a wireless network at a UE during a handover.
- **Delay jitter** is the time difference between the previous packet from the previous network and the next packet from the target network.

![Figure 5. Implementation Environment of SmSCTP.](image-url)
5.1. Signaling Cost

In analyzing the handover performance of SIP, mSCTP and SmSCTP, this work assumes the mobility pattern follows a uniform distribution and every handover event is independent. Table I lists the parameters used in this analysis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{req}$</td>
<td>Request cost</td>
</tr>
<tr>
<td>$C_{res}$</td>
<td>Response cost</td>
</tr>
<tr>
<td>$C_{hb}$</td>
<td>Heartbeat cost</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of handoff time</td>
</tr>
<tr>
<td>$\Psi_{SIP}$</td>
<td>Total signaling cost of SIP</td>
</tr>
<tr>
<td>$\Psi_{mSCTP}$</td>
<td>Total signaling cost of mSCTP</td>
</tr>
<tr>
<td>$\Psi_{SmSCTP}$</td>
<td>Total signaling cost of SmSCTP</td>
</tr>
</tbody>
</table>

Figures 6(a) and (b) illustrate the signaling costs of SIP and mSCTP, respectively. In SIP, a UE enables to send a re-INVITE message to demand the CN to perform the handover procedure. However, a conventional SIP re-INVITE procedure needs four requests and three responds between the UE and CN through the CSCF. Additionally, after handover completed, the UE executes the registration procedure to update the database of the CSCF. Therefore, the signaling cost of SIP can be obtained as follows.

$$\Psi_{SIP} \approx n[5CMR \times C_{req} + 4CMR \times C_{res}] = nCMR \times [5C_{req} + 4C_{res}], \quad (4)$$

where $0 < CMR < \infty$

In mSCTP, a UE is able to maintain the network status using heartbeat message. The heartbeat message is sent sequentially. If three heartbeat-reply message do not receive, the UE then performs the handover procedure to re-set primary connection. Note that the signaling message of mSCTP are only transmitted between the UE and CN. Moreover, to maintain global reachability, the UE needs to perform SIP registration procedure with the CSCF server. Hence, the signaling cost of mSCTP can be stated:

$$\Psi_{mSCTP} \approx n[7CMR \times C_{req} + CMR \times C_{res} + 3CMR \times C_{hb}] = nCMR \times [7C_{req} + C_{res} + 3C_{cost_{hb}}], \quad (5)$$

where $0 < CMR < \infty$

SmSCTP is based on the cross-layer design, and only uses two signaling messages for inserting new IP information and switching primary connection. In addition, the SSESSION message is able to register itself to the CSCF and CN. Therefore, the current registration procedure can be ignored. The signaling cost of SmSCTP can be stated as follows.
\[
\Psi_{SmSCTP} \approx \left[ 4CMR \ast C_{req} + 4CMR \ast C_{res} \right] + (n - 1) \ast \left[ 2CMR \ast C_{req} + 2CMR \ast C_{res} \right] \\
= 2CMR \ast (n + 1) \ast (C_{req} + C_{res})
\] (6)

where \(0 < CMR < \infty\)

Finally, the logarithms of 4, 5 and 6 are calculated. That is

\[
\Psi'_{SIP} = \log \Psi_{SIP}
\] (7)

\[
\Psi'_{mSCTP} = \log \Psi_{mSCTP}
\] (8)

\[
\Psi'_{SmSCTP} = \log \Psi_{SmSCTP}
\] (9)

Figures 7(a) and (b) plot the impact of CMR, which ranges at 0.5-1.5, on signaling cost. A high CMR value means that the UE moves frequently; otherwise, the UE stays on the same network for a long time. According to the CMR change value, however, the system load can
be determined. In this case, signaling cost of SmSCTP is significantly less than that of the SIP re-INVITE and mSCTP schemes because SmSCTP only uses four messages in each cycle-ANNEXIP, SSESSION and two 200 OK messages-to finish the handover and REGISTER procedures. However, the SIP re-INVITE scheme requires nine messages in each cycle and the mSCTP uses 11 messages in each cycle. When \( n > 1 \), SmSCTP only uses two signaling messages-one SSESSION and one 200 OK-to complete the handover procedure because the CN contained all IP information of the UE according to the ANNEXIP message. Therefore, the slope of SmSCTP is lower than that of SIP and mSCTP. Figure 7(c) shows implementation results versus average CMR value. The lines are the average mathematic results and points are the implementation results. Analytical results show that the proposed implementation outcomes are very close to the curve of the average CMR value. Both the mathematical and implementation results prove that the proposed scheme is correct.

Figure 7. Impact of call mobility ratio on signaling cost and real implementation. (a) \( CMR = 0.5 \), (b) \( CMR = 1.5 \), (c) real implementation results vs. Avg. CMR
5.2. Handover Latency

Table II lists parameters for measuring handover latency. For simplification, this work ignores transmission delay time between the GGSN and DHCP. The distance between the GGSN and UE equals that between the GGSN and CSCF.

Table II. List of the parameters for handover latency

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$</td>
<td>Delay time between UE and SGSN</td>
</tr>
<tr>
<td>$T_2$</td>
<td>Delay time between SGSN and GGSN</td>
</tr>
<tr>
<td>$T_3$</td>
<td>Delay time between GGSN and CSCF</td>
</tr>
<tr>
<td>$T_4$</td>
<td>Delay time between WAG and CSCF</td>
</tr>
<tr>
<td>$T_5$</td>
<td>Delay time between WAG and CSCF</td>
</tr>
<tr>
<td>$T_6$</td>
<td>Delay time between GGSN and CN</td>
</tr>
<tr>
<td>$T_{DHCP}$</td>
<td>Delay time for obtains IP address</td>
</tr>
<tr>
<td>$BW_{w_{UMTS}}$</td>
<td>Bandwidth of the UMTS access network</td>
</tr>
<tr>
<td>$BW_{w_{UMTS}}$</td>
<td>Bandwidth of the UMTS/HSDPA core network</td>
</tr>
<tr>
<td>$BW_{w_{PDN}}$</td>
<td>Bandwidth of the PDN</td>
</tr>
<tr>
<td>$S_{SIP}$</td>
<td>Average size of the SIP message</td>
</tr>
<tr>
<td>$S_{mSCTP}$</td>
<td>Average size of the mSCTP message</td>
</tr>
<tr>
<td>$S_{SmSCTP}$</td>
<td>Average size of the SmSCTP message</td>
</tr>
<tr>
<td>$L_{w_{UMTS}}$</td>
<td>Latency of the UMTS/HSDPA access network</td>
</tr>
<tr>
<td>$L_{PDN}$</td>
<td>Latency of the PDN network</td>
</tr>
</tbody>
</table>

Figure 8 depicts the message flows for delay analysis of the SIP re-INVITE and mSCTP schemes. In the SIP re-INVITE scheme (Fig. 8(a)), the UE waits for PDP context activation completion because SIP is a hard handover scheme. Let $D_{PDP}$ be the time needed by the PDP context activation procedure. Thus, $D_{PDP}$ is equal to $3T_1 + 3T_2 + T_{DHCP}$. Let $D_{SIP}$ be the delay time of the SIP re-INVITE procedure. The equation of delay time is

\[
D_{SIP} \approx nCMR(4T_1 + 4T_2 + 4T_3 + 3T_5 + D_{PDP})
= 4nCMR\left[ \left( \frac{S_{SIP}}{BW_{w_{UMTS}}} + L_{w_{UMTS}} \right) + 2\left( \frac{S_{SIP}}{BW_{w_{UMTS}}} + L_{w_{UMTS}} \right) + nCMR \right]
= 3 \left( \frac{S_{SIP}}{BW_{PDN}} + L_{PDN} \right) + D_{PDP}
\]

(10)

In Fig. 8(b), the $D_{PDP}$ needs not be included because the connection is continuing while handover occurs. However, the UE takes a long period of time to gather and determine the network status using mSCTP heartbeat message. Therefore, $D_{mSCTP}$ is

Copyright © 2000 John Wiley & Sons, Ltd.

Int. J. Commun. Syst. 2000; 00:1–6

Prepared using dcauth.cls
Figure 8. Message flow for delay analysis (a) SIP re-INVITE (b) mSCTP.

\[
D_{mSCTP} \approx CMR * n * [4T_1 + 4T_2 + 4T_6] = 4nCMR\left[\left(\frac{S_{mSCTP}}{BW_{wPMTS}} + L_{wL}\right) + \left(\frac{S_{mSCTP}}{BW_{wPMTS}} + L_{w}\right)\right] + 4n * CMR\left(\frac{S_{mSCTP}}{BW_{wPDP}} + L_{PDP}\right)
\]

Figure 9 depicts the proposed SmSCTP scheme for handover latency. According to the multi-homing concept, the primary connection works continually until it is broken. A UE has a sufficient time to establish its backup connection (e.g., a WLAN attachment) before the connection is broken. SmSCTP, on the other hand, uses the active handover mechanism. Hence, SmSCTP only uses one round trip time to complete the handover procedure. As with mSCTP, \(D_{PDP}\) can be ignored. Hence, the \(D_{SmSCTP}\) is defined as follows.
D_{SmSCTP} \approx CMR \times n \times (2T_1 + 2T_2 + 2T_3 + 2T_5) \quad (12)

Because T_3 = T_5,

\[ D_{SmSCTP} = 2nCMR\left[\left(\frac{S_{SmSCTP}}{BW_{wU,MTS}}\right) + L_{wl}\right] + \left(\frac{S_{SmSCTP}}{BW_{wU,MTS}}\right) + L_w \]

\[ + 4nCMR\left(\frac{S_{SmSCTP}}{BW_{wP,PDN}}\right) + L_{PDN} \quad (13) \]

Eqs. 10, 11, and 13 are estimated by applying to the logarithm function.

Figures 10(a), (b), and (c) plot the impact of CMR on delay time. We assume all moving directions of the UE are from the WLAN to the UMTS/HSDPA network. Consequently, the handover latency of SmSCTP is better than other two schemes because the proposed scheme is able to switch connection actively. In Fig. 10(a), if a user moves infrequently (n < 3), the delay time of SmSCTP is not clear. When n > 3, SmSCTP achieves short handoff time, if the CMR value exceeds 1 (Fig. 10(b)). Figure 10(c) shows the real delay time of SIP, mSCTP, and SmSCTP and average CMR. This scenario assumes the PDP context procedure lasts 2 seconds.

The line with circuit in Figure 10(c) is SmSCTP scheme. Generally, SmSCTP completes the handoff procedure using the round-trip time of two signaling messages. Thus, during each handoff, SmSCTP completes the procedure within approximately 200 milliseconds; conversely, SIP takes roughly 2-3 seconds, and the mSCTP needs 400-500 milliseconds. The average CMR value is the mobility characteristic of most humans. Implementation outcomes for SIP, mSCTP, and SmSCTP are very close to the average CMR curve, meaning that the proposed scheme is suitable for heterogeneous networks.

5.3. Packet Loss

To evaluate accurately the packet loss and delay jitter of SmSCTP, simulation is performed by extending the network simulator NS-2.27 [37]. For the SIP protocol, this work patches
Figure 10. Impact of call mobility ratio on delay time and real implementation. (a) $CMR = 0.5$, (b) $CMR = 1.5$, (c) real implementation results vs. Avg. $CMR$

the simulator using the SIP patch [38]. In our simulation, a UE has two wireless access interfaces-WLAN and UMTS/HSDPA. The bandwidth of WLAN is 11Mbps and the latency is 10ms; the bandwidth of the UMTS/HSDPA network is 384Kbps and the latency is 70ms. The simulation used the SIP module and mSCTP module to simulate the handoff procedures in a heterogeneous network environment. Moreover, the Constant bit rate (CBR) method is used in NS-2 simulator to generate the UDP traffic. The payload of CBR traffic is 1Kbytes and the packet arrival rate is set up to 125 packets per second.

Figure 11 shows the packet loss for three handover schemes, e.g. the SmSCTP, SIP re-INVITE and mSCTP schemes. The movement direction of the UE is shown in top of figures. In Fig. 11(a), the SIP re-INVITE scheme is blocked for acquiring a new IP address because of hard handover scheme. Therefore, packets are still sent to previous location until the handover completes. On the other hand, notably, the transmission rate of the UMTS/HSDPA network is
less than that of WLAN. As a consequence, the handoff delay from UMTS/HSDPA to WLAB is more than another one. The average number of SIP packet losses is roughly 330 packets. However, the mSCTP and SmSCTP schemes are multihoming approaches. That is, the UE can acquire two or more IP addresses from both networks, simultaneously. Hence, packets can be sent to other backup location when the primary location is down. For mSCTP or SmSCTP, the effect of packet loss is negligible. Average packet loss for mSCTP and SmSCTP are roughly 125 and 75 packets, respectively.

Figure 11. Packet loss. (a) WLAN–> UMTS (b) UMTS–> WLAN

5.4. Delay Jitter

Figure 12 shows the delay jitter for the SmSCTP, SIP re-INVITE and mSCTP schemes. Figures 12(a) and (b) show the moving direction from a WLAN to a UMTS and from a UMTS to a WLAN, respectively. In Fig. 12(a), the SIP re-INVITE scheme takes approximately 3100-3200 milliseconds for handover processing because it has to complete the PDP context attachment procedure; mSCTP and SmSCTP take approximately 1000 milliseconds and 600 milliseconds, respectively, because mSCTP and SmSCTP use multihoming to process the PDP context attachment procedure in parallel.

In Fig. 12(b) experimental results demonstrate that the delay jitter of the SIP re-INVITE takes approximately 4000-4100 milliseconds because the SIP re-INVITE scheme conducts the PDP context de-attachment procedure. Thus, the SIP re-INVITE scheme spends more time to complete the handoff procedure than mSCTP and SmSCTP. The experimental results for the mSCTP and SmSCTP are the same as Fig. 12(a) because the PDP context de-attachment procedure can be processed with transmission data in parallel through the WLAN network.

6. Conclusions

The network heterogeneity is a intrinsic property that allows a UE to access the Internet services anytime, anywhere due to the convergence of different wireless technologies. In such
Figure 12. Delay jitter. (a) WLAN → UMTS (b) UMTS → WLAN.

networks, the UE acquires a robust communication without care it is at indoor and outdoor. This imposes a challenge that requires a novel mobility scheme to accommodate the evolving complexities in heterogeneous wireless networks. In this work, we have represented a novel cross-layer session mobility in WLAN/3.5G heterogeneous wireless networks. The proposed scheme consists a better lifetime based handover policy and SIP-based mSCTP scheme, called SmSCTP. In handover policy, the UE selects a access point to achieve better goodput. In addition, SmSCTP provides a seamless handover scheme which is integrated by application-layer and transport-layer. Two novel SIP signaling messages are utilized for SmSCTP. With these SIP messages, the UE establishes quickly a backup path before handoff and the CSCF server can acquire the new IP address information for the UE. Mathematical analyses verify that the proposed SmSCTP scheme offers a better latency and signaling cost than the SIP re-INVITE scheme and mSCTP. Finally, implementation results demonstrate that SmSCTP achieves performance improvements for handoff delay and delay jitter.

Our future work will take account to extend our scheme for different wireless networks, such as Long term Evolution (LTE) and Vehicular Ad hoc Networks (VANETs). In such networks, a mobile device can move with high velocity. Hence, the mobile device needs a fast and seamless mobility management scheme to maintain connectivity than the conventional WLAN/3.5G heterogeneous networks. Moreover, the future work also considers the characteristics of the network and client. To address this problem, we will develop a novel mobility management compacted with current protocols for such wireless networks.

7. Remark

This section compares the SIP re-INVITE and mSCTP schemes with SmSCTP. Table III shows the comparisons of these three schemes. The major characteristics of each method are compared, and formal definitions are shown as following.

The "Handover Type" can point out that a UE can set up more than one connection at
Table III. Comparison of the SIP re-INVITE, mSCTP, and SmSCTP schemes

<table>
<thead>
<tr>
<th>Properties</th>
<th>SIP re-INVITE</th>
<th>mSCTP</th>
<th>SmSCTP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handover Type</td>
<td>Hard</td>
<td>Soft</td>
<td>Soft</td>
</tr>
<tr>
<td>Protocol Layer</td>
<td>Layer-7</td>
<td>Layer-4</td>
<td>Layer-7 and -4</td>
</tr>
<tr>
<td>Handover Latency</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Multi-streaming</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Multi-homing</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Scalability</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Flexibility</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
</tr>
</tbody>
</table>

the same time or only one. The SIP re-INVITE scheme is a hard handover scheme. Thus, the connection of the SIP re-INVITE scheme should be broken before setting up a new one. The mSCTP and SmSCTP schemes are based on soft handover and enable to build up multi-connection before switching the wireless access point. In "protocol type", the SmSCTP scheme is a cross-layer method, and provides a handover policy and low signaling cost than other two schemes. The "handover latency" can responds the comfortability on call for a user. SmSCTP achieves better handover latency than other schemes. In "multi-homing", mSCTP and SmSCTP are able to set up more than one backup connection through different wireless networks. In addition, the SIP re-INVITE scheme needs to send another INVITE message to the CN to switch the wireless network. Before the INVITE message arrivals, the packet is lost.

In "multi-streaming", the voice/video packets of SmSCTP are only transmitted using primary path. Therefore, SmSCTP does not support multi-streaming and the transmission efficiency of mSCTP is better than SmSCTP. In "scalability", the SIP re-INVITE and SmSCTP schemes use infrastructure to maintain all UEs. The SIP messages must be sent to the CSCF server. Therefore, both of these schemes cause a bottleneck in the CSCF server. The mSCTP scheme is an end-to-end protocol. The control messages of mSCTP are sent to another endpoint directly. The mSCTP can thus acquire a better scalability than the SIP re-INVITE and SmSCTP schemes. In "flexibility", the SIP re-INVITE and SmSCTP schemes are belonged to layer-7 protocol. The UE can be easy to add or modify the protocol using software update mechanism. On the other hand, the user of mSCTP needs to modify the operation system to support a better performance or a new function.

Acknowledgment

This research was supported by the National Science Council of the R.O.C. under grants NSC-97-2219-E-305-001 and NSC-98-2219-E-305-001. This research is an extension of results present to IEEE Wireless Communications and Networking Conference, 2007(WCNC’07).

REFERENCES


5. 3GPP. Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2. 3GPP Std., May 2006.

6. 3GPP system to Wireless Local Area Network (WLAN) interworking; System description (Release 6), 3GPP Std., Rev. TS 3G TS 23.234 V6.0.0 (2004-03); March 2004.

7. 3GPP. 3GPP system to Wireless Local Area Network (WLAN) interworking; System description (Release 7), 3GPP Std., Rev. TS. 3G TS 23.234 V7.1.0 (2006-03); March 2006.


23. Snoeren A-C, Balakrishnan H. An End-to-End Approach to Host Mobility, in Proceedings of the 6th annual international conference on Mobile computing and networking (MobiCom’00); Aug. 06-11, 2000; 155-166.


29. Farrokh A, Krishnamurthy V. Opportunistic scheduling for streaming multimedia users in high-speed
AUTHORS’ BIOGRAPHIES

Kuan-Lin Chiu received the B.S. and M.S. degrees from Providence University, Chung-chi Rd., Taichung 43301, Taiwan, R.O.C., in 2002 and 2005, respectively. He is currently working toward the Ph.D. degree in computer science and information engineering, at the National Chung-Cheng University (CCU), Chia-Yi, Taiwan, R.O.C. He has received the CSIE Ph.D candidate from the National Chung Cheng University. His current research interests include Vehicular Ad hoc Network (VANET), heterogeneous wireless networks, cross-layer design, mobile WiMAX networks, computer telephony, and mobility management.

Yuh-Shyan Chen received the B.S. degree in Computer Science from Tamkang University, Taiwan, R. O. C., in June 1988 and the M.S. and Ph.D. degrees in Computer Science and Information Engineering from the National Central University, Taiwan, R. O. C., in June 1991 and January 1996, respectively. He joined the faculty of Department of Computer Science and Information Engineering at Chung-Hua University, Taiwan, R. O. C., as an associate professor in February 1996. He joined the Department of Statistical, National Taipei University in August 2000, and joined the Department of Computer Science and Information Engineering, National Chung Cheng University in August 2002. Since 2006, he has been a Professor at the Department of Computer Science and Information Engineering, National Taipei University, Taiwan. Prof. Chen is now serving as Vice Chair of Task Force on "Telecommunications" of Intelligent Systems Applications Technical Committee, IEEE Computational Intelligence Society. Prof. Chen served as Editor-in-Chief of International Journal of Ad Hoc and Ubiquitous Computing (SCI), Regional Editor (Asia and Pacific) of IET Communications (SCI), Editorial Board of Telecommunication System Journal (SCI), EURASIP Journal on Wireless Communications and Networking (SCI), International Journal of Communication Systems (SCI), and Mobile Information Systems (SCI). He also serves as Guest Editor of ACM/Springer Mobile Networks and Applications (MONET), and Wireless Communications and Mobile Computing, The Computer Journal, and Wireless Personal Communications. His paper wins the 2001 IEEE 15th ICOIN-15 Best Paper Award. Prof. Chen was a recipient of the 2005 Young Scholar Research Award, National Chung Cheng University, R.O.C.. His recent research topics include wireless communications, mobile computing, and next-generation personal communication system. Dr. Chen is a senior member of the IEEE Communication Society and Phi Tau Phi Society.

Ren-Hung Hwang received his B.S. degree in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 1985, and the M.S. and Ph.D. degrees in computer science from University of Massachusetts, Amherst, Massachusetts, USA, in 1989 and 1993, respectively. He joined the Department of Computer Science and Information Engineering, National Chung Cheng University, Chia-Yi, Taiwan, in 1993, where he is now a distinguished professor and the director of Ching-Jiang Learning Center. Since 2010, he is also an adjunct professor of the Department of Computer Science and Information Engineering, National Sun Yat Sen University, Kao-Hsiung, Taiwan. His research interests include ubiquitous computing, peer-to-peer applications, wireless networks, and e-Learning.