

Moving toward end-to-end support for handoffs across heterogeneous telephony systems on dual-mode mobile devices

Hung-Yun Hsieh^{a,b,*}, Chung-Wei Li^b, Shuo-Wei Liao^b, Yu-Wen Chen^b,
Tsung-Lin Tsai^b, Hsiao-Pu Lin^b

^a Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan

^b Graduate Institute of Communication Engineering, National Taiwan University, Taipei 106, Taiwan

Available online 21 April 2007

Abstract

More and more handsets manufactured today are being equipped with WLAN modules, allowing users to access VoIP service whenever they are within the coverage of WLAN service area. One of the greatest challenges for such dual-mode (say GSM and WiFi) handsets to support ubiquitous communication is the ability to maintain session continuity so a phone call established through one mode can be transferred to the other mode without user intervention. While related work has investigated the problem of supporting vertical handoffs for data communications across heterogeneous wireless data networks (e.g. GPRS and WiFi), these approaches cannot be directly used for vertical handoffs across heterogeneous telephony systems that involve both *circuit-switched* and *packet-switched* systems. In this paper, we investigate the problem of supporting vertical handoffs for voice communications across circuit-switched and packet-switched telephony systems on dual-mode mobile devices. We first contend that an end-to-end solution has advantages over infrastructure-based solutions due to the unique characteristics of the problem considered. We then propose an end-to-end approach that builds atop existing cellular network and VoIP infrastructure without introducing new entities for supporting vertical handoffs on dual-mode mobile devices. The proposed approach is based on the Session Initiation Protocol (SIP) and requires modifications only on end devices. We conduct testbed experiments using a GSM-WiFi dual-mode PDA and a SIP soft phone, and evaluate the quality of speech on the dual-mode PDA when the call is migrated from WiFi to GSM. Evaluation results show that the proposed approach allows soft handoffs and is a promising solution for seamless vertical handoffs on dual-mode mobile devices.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Heterogeneous wireless networks; Multi-homed mobile device; Dual-mode handset; Vertical handoff; VoIP; SIP

1. Introduction

VoIP over WLAN (VoWLAN) has attracted a lot of attention recently due to the prevalence of the VoIP service and the IEEE 802.11-based WLAN technology (WiFi). Compared with other wireless access technologies, WLANs can be deployed rather quickly and offer much higher data rate with substantially lower cost. Placing VoIP calls from inside these WiFi hot spots hence enables untethered voice

communications without incurring the cost required by existing mobile cellular telephony systems such as GSM, PCS, IS-95, and 3G.

Despite the potential benefits of the VoWLAN technology, the fact that WLANs typically have only spotty coverage limits the applicability of VoWLAN-enabled handsets. A VoWLAN call can be placed or received only when the user is within the reach of WLANs. When moving outside the coverage of WiFi hot spots, for example, a mobile user may need to resort to his/her GSM handset for assurance of voice communications. While more and more cities worldwide have undertaken plans to build large-scale WiFi networks or even WiMAX networks with the goal of enabling city-wide wireless Internet access, the coverage

* Corresponding author. Address: Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan. Tel.: +886 2 33663666.

E-mail address: hyhsieh@cc.ee.ntu.edu.tw (H.-Y. Hsieh).

still cannot parallel that provided by existing cellular telephony systems, especially in suburban or rural areas.

To take advantage of the *opportunistic* nature of the VoWLAN service without the need to bring two handsets, the concept of dual-mode handsets has emerged and come to the limelight. Many handset manufacturers including Motorola, Nokia, and NEC have released GSM–WiFi dual-mode handsets that can be used for making GSM or VoWLAN calls. Several EDGE–WiFi, 3G–WiFi, and HSDPA–WiFi dual-mode handsets and PDA phones have also made their debut recently [1–3]. Such dual-mode handsets¹ by design can be used as either GSM or VoWLAN handsets at different times and places depending on service availability and user preference. However, a key functionality that truly differentiates *one* dual-mode handset from *two* single-mode handsets is the ability to allow *seamless migration of a phone call from one telephony system to the other*. Take the GSM–WiFi dual-mode handset for example. A call connected through the WiFi mode (VoWLAN call) should be migrated to the GSM mode (GSM call) when the user moves beyond the reach of WiFi service. If handoffs can take place automatically (subject to user preference) without disrupting the quality of ongoing voice communications, the benefits of dual-mode handsets can be truly exploited.

Related work has investigated the problem of vertical handoffs across heterogeneous wireless overlay networks using techniques such as packet buffering and forwarding, double-casting, path re-routing, and packet striping for achieving fast and seamless handoffs [4–7]. These solutions, however, cannot be applied to solve the problem faced by dual-mode handsets. GSM, like other mobile cellular and public switched telephone networks, is a *circuit-switched telephony system* that establishes the call using dedicated “circuits” without following the *store-and-forward* paradigm in packet-switched data networks. In circuit-switching networks, a circuit that is dedicated to one call in progress cannot be used by other callers until the circuit is released. Therefore, while it is possible to multiplex packets from different wireless data networks (e.g. through packet buffering and forwarding at the base station) for ensuring seamless call migration during handoffs, it is non-trivial to multiplex one circuit-switched call with one packet-switched call using only packet-based techniques. Moreover, circuit-switched networks have adopted the “*smart network, dumb terminal*” paradigm opposite to that

in the Internet. Hence it is difficult, if not impossible, to control circuit-switched voice effectively on end devices – as can be done in packet-switched networks (e.g. through manipulation of the packet header for re-routing). While it is possible to leverage infrastructure support for *interworking* the two telephony systems and performing conversion between circuit-switched and packet-switched voice inside the network [8–10], it might not be desirable for the dual-mode user due to the *extra cost and overhead incurred by directing VoWLAN traffic to the GSM core network*. Moreover, these infrastructure-based approaches are typically tailored to the specificity of wireless systems in consideration, and hence they cannot be easily extended to interwork new wireless systems such as WiMAX.

In this paper, we investigate the problem of end-to-end support for voice communications across heterogeneous telephony systems on dual-mode mobile devices that include handsets, smart phones, PDAs, and other portable devices equipped with both wireless NICs and/or modems. Specifically, we consider a scenario that involves voice communications between a dual-mode mobile device and a remote VoIP client. The remote VoIP client can be *a dual-mode mobile device, a wired or wireless softphone, or a VoIP gateway* installed in the enterprise, campus, or residential building. The goal is to allow the dual-mode device to switch transparently between WiFi and GSM modes during a phone call with the remote VoIP client. While there are many important issues such as AAA (authentication, authorization, and auditing) and mobility management that need to be addressed for such dual-mode mobile devices to switch from one system to the other, in this paper we focus on the issue of handoff management. We assume that the dual-mode device can make phone calls with the remote client through either mode of communication already, and we focus on the procedures that should be designed to ensure session continuity. Motivated by the high mobile handset replacement rate and the fact that more and more sophisticated functionalities are being added to the handsets, we target a solution that can be implemented on the end devices without requiring new network entities for vertical handoffs to be deployed.² We believe that future wireless networks can benefit from more intelligent mobile handsets to offload the burden of the network infrastructure for services such as vertical handoffs.

Toward this goal, we propose an end-to-end solution for vertical handoffs across heterogeneous telephony systems based on the Session Initiation Protocol (SIP), the standards developed by IETF for initiating, modifying, and terminating multimedia sessions including voice communications [11]. SIP is an end-to-end protocol that

¹ In such dual-mode handsets, each mode may still involve a combination of wireless technologies in different frequency bands (e.g. GSM 900/1800/1900 for the GSM mode, and 802.11 a/b/g for the WiFi mode), but a distinguishing feature from conventional “multi-band” handsets is the ability to support voice communications *simultaneously* through *heterogeneous* telephony technologies involving circuit-switched cellular telephony and packet-switched Internet telephony. While it is possible that the GSM mode may be replaced by 3G, and the WiFi mode may be replaced by WiMAX in the future, for ease of presentation, we collectively refer to these devices as GSM–WiFi dual-mode devices hereafter due to their prevalence.

² We note that there are already interworking entities such as SIP–PSTN gateways in existing VoIP infrastructure for bridging the two telephony systems. However, the mere existence of such gateways does not ensure *session continuity* for dual-mode mobile devices. Hence, new approaches that build atop existing infrastructure still need to be designed for the target problem.

works independently of the type of session that is being established, and it has been designed to interwork transparently with PSTN signaling including SS7 ISUP (ISDN User Part) [12]. Unlike SS7, however, SIP requires a very simple core network with intelligence distributed to communicating endpoints. We start by investigating different approaches based on SIP that have been proposed in different contexts for supporting mobility including using RE-INVITE for mid-call mobility [13] and using REFER for call transfer [14]. We identify the problems of these approaches when used directly in the target environment and propose an approach that essentially creates a 3-party conferencing call between dual-modes and the remote peer during handoffs to ensure “make-before-break” soft handoffs. We present details of the proposed solution in Section 4. Testbed evaluation using a dual-mode GSM–WiFi PDA and a SIP softphone shows encouraging results and motivates further research along this direction.

The rest of this paper is organized as follows. Section 2 discusses related work that adopts infrastructure-based approaches for vertical handoffs. Section 3 introduces SIP, and discusses existing solutions based on SIP mobility support. Section 4 proposes the new vertical handoff solution and discusses its operations in different network scenarios. Section 5 presents testbed results, and Section 6 discusses some issues and future directions. Finally, Section 7 concludes the paper.

2. Related work

While network heterogeneity has attracted a lot of attention recently, the problem considered in this paper is quite different from that in related work due to the nature of the telephony systems involved. Related work on seamless handoffs between heterogeneous wireless networks has focused on data access using IP as the underlying platform [4–7,10]. This paper, on the other hand, considers *voice communications across circuit-switched (GSM) and packet-switched (WLAN) networks*. Note that GSM uses a very different protocol stack from the WLAN in terms of both signaling and data transport. For example, GSM circuit-switched voice, after proper digitalization, coding, and interleaving, is fitted into the TDMA frames for transmissions, while WLAN packet-switched voice undergoes an entirely different protocol stack including codec selection, RTP packetization, UDP multiplexing, IP switching, and 802.11 MAC framing. Hence, conventional approaches that rely on all-IP solutions (e.g. mobile IP or TCP-based approaches) cannot be used in this context. While it is possible to support voice communications through GPRS packet data access (VoIP over GPRS), it is not practical due to the apparent lack of competitiveness in cost and quality compared to the GSM circuit-switched voice service.

As we mentioned in Section 1, many handset manufacturers have already released dual-mode handsets. However, not all dual-mode handsets manufactured today are equipped with the required functionality to support seam-

less handoffs between two heterogeneous telephony systems. Many handsets require the user to select the mode of communications manually, where handoffs from one mode to the other are possible only through *manual redialing* using the new mode. Some handsets can switch between the two modes automatically (thus freeing the user from redialing) but the user may experience session discontinuity during an active call due to *hard handoffs*. Several dual-mode handsets *do* allow soft handoffs between the two modes without user intervention. However, they have been designed for use with customized support from the GSM or WLAN infrastructure. For example, Motorola’s CN620 dual-mode handsets rely on the MAP architecture developed jointly by Motorola, Avaya, and Proxim to provide seamless handoffs in target environments [15]. Azaire Networks Inc. develops an “IP Converged Network Platform (IP-CNP)” using a 3GPP-WLAN convergence gateway, and teams up with several telecom service providers for deployment. VeriSign’s “Wireless IP Connect Service” and Calypso Wireless’ ASNAP solution are yet another two examples.³ Apparently, these proprietary solutions are tailored to specific dual-mode handsets and service providers. Hence, such dual-mode handsets cannot be used in other systems operated by different service and equipment providers. The interoperability of different incompatible systems could eventually become an issue hurdling the popularity of such dual-mode handsets.

Recently, an open standard called Unlicensed Mobile Access (UMA) has been developed by the UMA Consortium. It is later adopted by the 3rd Generation Partnership Project (3GPP) as the specification for *Generic Access (GA) to the A/Gb interfaces* in the public land mobile network (PLMN, such as the GSM core network) [8]. The goal is to extend GSM mobile service over an IP network (called Generic Access Network or GAN) so mobile stations can obtain services from the GSM core network through IP access rather than through the traditional GERAN (GSM/EDGE Radio Access Network) radio interface. As shown in Fig. 1, a new network element called GAN controller (GANC) is introduced in the GAN architecture. GANC appears to the core network as a GERAN BSS (Base Station Subsystem), where it interworks with the core network through the A (circuit-switched service) and Gb (packet-switched service) interfaces, and with the mobile station (MS) through the Up (IP data transport) interface. To bridge circuit-switched service in the PLMN and packet-switched data in the GAN, GANC performs through the A-interface functionalities including re-framing from RTP packets to A-interface frames, transcoding voice to/from the MS to PCM voice from/to the MSC, and establishment, administration and release of control/user plane bearers between the MS and the core network.

³ Detailed information on individual products and solutions from Azaire Networks (<http://www.azairenet.com>), VeriSign (<http://www.verisign.com/>), and Calypso Wireless (<http://www.calypsowireless.com>) is available online.

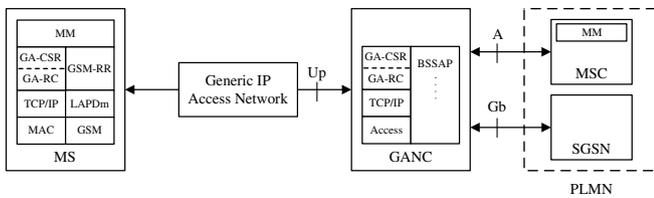


Fig. 1. GAN functional architecture (with CS domain control plane).

In addition to changes at the GANC, the GAN architecture introduces two sublayers atop the IP transport layer at the MS called GA-CSR (Circuit-Switched Resources) and GA-RC (Resource Control) to peer with the GSM-RR (Radio Resources) sublayer of the GSM protocol stack as shown in Fig. 1. GA-RC provides functionalities including discovery, registration, and keep-alive with the GANC. GA-CSR sits atop GA-RC and provides functionalities including setup of bearer for CS traffic between the MS and GANC, and handover support between the GERAN and GAN. Note that in the GAN architecture, the GANC terminates the packet-switched voice toward the MS, and hence the IP access network (e.g. WLAN) appears as another radio access network (e.g. GERAN) to the core network. Therefore, mobility management (MM) at the MSC including handoffs can be performed as usual. Several cellular telephony service providers and handset manufacturers have announced their plans to support the GAN architecture.

The advantage of the GAN architecture is that it integrates heterogeneous wireless data networks into the PLMN for providing a unified access to the mobile device, and it is possible to use legacy devices such as POTS (Plain Old Telephone Service) phones for communications with the dual-mode device. However, there is still motivation for dual-mode users to consider *solutions orthogonal to GAN*. For one, since all traffic from the WLAN traverses through the GSM system, the dual-mode user will be charged for using service provided by the GSM core network (note that the packet-switched voice from MS is terminated at the GANC) when placing VoWLAN calls with the purpose to *save the cost incurred in using the GSM phone service*. Also, WLAN traffic has to flow through the GSM core network, incurring additional latency and introducing bottleneck to voice communications between the dual-mode user and the remote peer. The problem will be more serious if GANC's are deployed without sufficient density. Finally, with the emergence of new wireless data technology such as WiMAX (IEEE 802.16), it is possible that there will be GSM–WiMAX or 3G–WiMAX handsets in the future. It remains to be seen whether infrastructure-based approaches such as GAN can be easily adapted to support new wireless systems.

3. SIP and mobility support

In this section, we first present a brief overview of SIP, an end-to-end application-layer signaling protocol that

works independently of the type of session that is being established, and then discuss why existing solutions for SIP mobility are not applicable. In Section 4, we propose a new vertical handoff solution based on SIP.

3.1. Session Initiation Protocol

The Session Initiation Protocol (SIP) is an application layer signaling protocol developed by IETF for creating, modifying, and terminating sessions with one or more participants [11]. A session in SIP is a collection of participants and the media streams (including audio, video, and text) between them for the purposes of communication. As a signaling protocol, SIP works in concert with other existing Internet protocols such as SDP (Session Description Protocol), RTP (Real-time Transport Protocol), and RTCP (Real-time Control Protocol) for control and delivery of multimedia data. While SIP works independently of the underlying transport protocols and the type of session that is being established, it has been designed to support traditional telephony services present in the public-switched telephone network (PSTN) such as call forwarding, call transfer, and 3-way conferencing.

SIP is based on an HTTP-like request/response transaction model, where the SIP end system (user agent or UA) that generates requests is called a user agent client, and the one that responds to requests is called a user agent server. A SIP transaction consists of a request from the client to invoke a particular method (function) on the server, and at least one response triggered by the request. In the original SIP specification [11], six methods are defined: REGISTER for registering contact information, INVITE, ACK, and CANCEL for setting up sessions, BYE for terminating sessions, and OPTIONS for querying servers about their capabilities. Several other methods such as INFO, REFER, and NOTIFY have also been defined in other RFCs [14,16] to extend the functionality of SIP for implementing different telephony services.

SIP is an end-to-end protocol that does not rely on any SIP-specific servers to create, modify, or terminate sessions between end systems as long as requests and responses can be delivered to target user agents. To facilitate directory lookup and location service, however, several SIP servers such as registration servers, redirect servers, and proxy servers are often introduced in SIP networks. Specifically, proxy servers help route requests to the user's current location, registration servers allow users to upload their current locations for use by proxy servers, and redirect servers help locate the user by providing alternative locations where the user can be reached. It has to be noted that while SIP proxies may help forward or respond to requests before the session is setup, they are not required once the end systems learn of the addresses of remote entities. More importantly, a SIP proxy does not need to understand a SIP request in order to forward it. A SIP proxy receiving a request with an unknown method will proxy the request without making any changes to the header. Such a design along with the

fact that the intelligence in a SIP network is located in end systems allows *new services to be created and deployed by changing only end systems without any changes in the SIP network*. This is very different from the signaling system in the PSTN where the intelligence lies in the core of the network and new services cannot be easily supported without upgrading the entire infrastructure.

To allow phone calls to be made between SIP phones and conventional PSTN phones, SIP providers have introduced gateways to interwork the two networks. A gateway works by acting as a *user agent* for the SIP network and a *terminating switch* for the PSTN. It typically consists of three components (collocated or distributed): signaling gateway for receiving signaling on the PSTN side and encapsulating it over IP for routing (and vice versa), media gateway for terminating a PCM trunk on the PSTN side and bridging it to packetized bit streams through RTP payloads (and vice versa), and media gateway controller for controlling the media gateway and converting between the PSTN signaling protocol and the SIP. Note that while a gateway may interface with different signaling protocols such as CAS (Channel Associated Signaling) and ISUP depending the nature of the PSTN call (e.g. from PBX trunk or mobile phone networks), it still *appears to the SIP network as a SIP user agent*. A call from the SIP phone to the PSTN phone can be made through proper configuration at the gateway, whereas a call from the PSTN phone can be routed to the SIP phone if the latter is given a global telephony number as that provided by E-NUM. In addition to bridging calls between PSTN and SIP networks, it is possible to use the SIP network as the backbone network between two PSTN phones. Interested readers are referred to [12] for technical details on the SIP-T framework for PSTN–SIP–PSTN interworking.

3.2. RE-INVITE for SIP mobility

In a wireless environment, it is possible that the IP address of the participant might change due to mobility (e.g. changes of the point of attachment). To keep the VoIP session after the participant acquires a new IP address, it is necessary to “modify” the session to reflect the new peer relationship and ensure proper multimedia delivery. In SIP, end systems can send an INVITE request within an *established session* (known as RE-INVITE) for modifying session parameters including changing addresses or ports, adding media streams, and deleting media streams. Related work has proposed mechanisms to allow mid-call mobility (handoffs during a call) by using the INVITE request during the call for modifying the IP address [13].

While SIP mobility based on the RE-INVITE method can tackle the problem of address change during an active session, it can not be used directly for handoffs between the GSM and WiFi modes on a dual-mode mobile device. The reason is that RE-INVITE only allows for change of session parameters between two end systems with an *established session*, but handoffs from WiFi to GSM modes

require establishment of a new session involving the third user agent – the PSTN gateway (or SIP–GSM gateway) – for connecting the GSM mode and bridging circuit-switched and packet-switched voice streams. Even if the IP address of the PSTN gateway is known *a priori* to the dual-mode device, one cannot simply send a RE-INVITE request to inform the remote peer of the new contact address of the dual-mode device (for handoffs from the WiFi mode to the GSM mode). The states on the PSTN gateway for the call between the GSM mode and the remote peer need to be established anew, so the voice can be redirected to the new path (session setup includes negotiation of media parameters such as the audio codec to use). Hence the SIP RE-INVITE method cannot be used “as-is” to solve the target problem.

3.3. REFER for SIP mobility

If we consider the GSM and WiFi modes on the dual-mode device as two independent users that might engage in communications with the remote peer, then handoffs between the two modes can be considered as *call transfer* from one user to the other. Therefore, another direction toward solving the target problem is through transfer of the two calls. The REFER method indicates that the recipient should contact a third party using the contact information provided in the request [14]. It has been proposed to extend the ability of SIP for supporting enhanced telephony services available in the PSTN, in particular the call transfer and third-party call control services. In the simplest form of call transfer called *blind (or unattended) transfer*, the transferor provides the transfer target’s contact (in the Refer-To header field) to the transferee, and then terminates the existing session with the transferee as soon as the latter acknowledges the request (without waiting for the outcome of call transfer). *Attended transfer*, on the other hand, involves setting up a session between the transferor and the transfer target before the transfer (from the transferee to the transfer target) actually takes place. The session between the transferor and the transferee is placed on hold when the new session is being established, and is terminated only after the the transfer is complete [17]. In this way, the transfer target can be notified of the impending transfer, while the transferor can be notified of the transfer progress and potentially assist the transferee with recovery from transfer failure.

Since the REFER method requires a new session be setup to which the old session is transferred, it does not suffer from the same problem in the RE-INVITE request as mentioned in Section 3.2. Note, however, that both the unattended and attended transfers adopted in conventional call transfer services still cannot be used “as-is” for addressing the target problem. In unattended transfer, the session between the old mode (transferee) and the remote peer is terminated before the session between the new mode (transfer target) and the remote peer is established. On the other hand, attended transfer involves con-

sultation hold that places the existing conversation temporarily on hold when the new call is being established. Such a “break-before-make” hard handoff cannot ensure seamless handoff especially when the call setup delay between the remote peer and the new mode is large. As we will show in Section 5, the call setup time between GSM and the remote SIP client is rather long (to the order of 10 s), and hence if *vanilla call transfer* is used for vertical handoffs from WiFi to GSM modes, then the disruption due to temporal discontinuity will be noticeable and unacceptable to the user. A desirable solution hence should ideally address this problem for achieving seamless vertical handoffs. We discuss in the following that without resorting to new methods how SIP can be used as the basis of the proposed solution.

4. The proposed solution

In this section, we first present the required changes at the end devices for supporting end-to-end handoffs without modifying existing infrastructure, and then we present detailed operations of the proposed handoff procedures when the dual-mode mobile device migrates from WiFi to GSM service, and vice versa.

4.1. Functional blocks

The proposed vertical handoff solution aims to ensure “make-before-break” soft handoffs while reusing the VoIP user agent and GSM phone modules on the dual-mode mobile device. Fig. 2 shows the proposed changes to the dual-mode device for achieving this goal.

The *VoIP UA* and the *GSM Phone* modules are used by the dual-mode device to make and receive calls through the WiFi and GSM modes respectively. The *Handoff Control* module is the core for handoff management. It controls when handoffs should be triggered through feedback of the *WiFi Monitor* module. The *Handoff Control* module also maintains states of the communications in progress including the identities (e.g. calling numbers) of the *VoIP*

UA and *GSM Phone* modules, and the policies (elaborated in Section 4.2) for performing vertical handoffs. It interacts with the *VoIP UA* and *GSM Phone* modules through commands such as *Call*, *Hangup*, and *Answer* for transparent call migration between the two communication modes. *Handoff Control* modules at both ends of the communications exchange information through *Handoff Trigger* and *Handoff Complete* commands and acknowledgements for initiating and finalizing the handoff process on end devices. It has to be noted that if the remote peer is a single-mode device (i.e. a SIP softphone), then one of the *VoIP UA* and *GSM Phone* modules will not be present in the device functional blocks, but the *Handoff Control* module is still present to facilitate the progress of the handoff process at the other end.

In the following, we consider scenarios where one end is a dual-mode device, and the other end is a SIP softphone (referred to as the remote peer). We discuss how these functional blocks are put together to perform the handoff process when the dual-mode device handoffs between WiFi and GSM service.

4.2. Handoff from WiFi to GSM

We consider the first scenario where the dual-mode device is initially in a call with the remote peer using the WiFi mode. During the call the dual-mode device may move outside the coverage of the WiFi service, and needs to migrate the call to the GSM mode to maintain session continuity. We assume that the remote peer is also a VoIP user agent as indicated by “Step 1” in Fig. 3. Section 6.1 discusses the scenario when the remote peer is a GSM phone.

During the call session, the *WiFi Monitor* module periodically measures the quality of the WLAN link and decides whether a handoff to the GSM mode should be performed. The link quality can be determined based on the received signal strength measurements or RTCP sender/receiver reports on packet loss rates, delays, and delay jitters. It is also possible to make handoff decision based on measurements of the speech quality using the chosen objective metrics [18]. If a handoff to the GSM mode is decided, the *Handoff Control* module sends a *Handoff Trigger* command to the remote peer to initiate handoff (as indicated by “Step 2” in Fig. 3). The *Handoff Trigger* command includes in the message the call state of the VoWLAN call (e.g. the call-ID in the SIP message header that uniquely identifies the call) and the GSM phone number so the remote peer can identify the dual-mode device and match the command

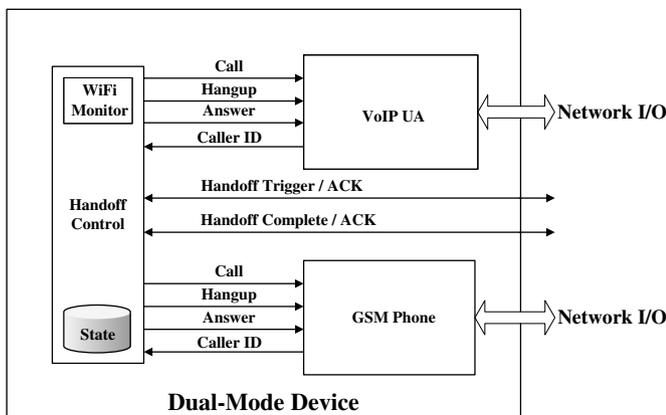


Fig. 2. Dual-mode device functional blocks.

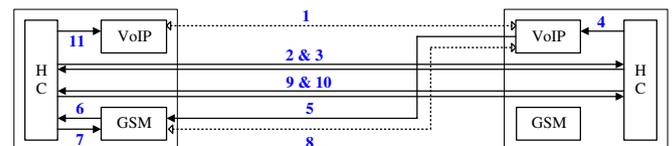


Fig. 3. Vertical handoff from WiFi to GSM (remote peer makes the second call).

to the existing call session. Upon receiving the handoff trigger, the *Handoff Control* module on the remote peer responds with *Handoff Trigger ACK* (“Step 3”), and then directs its *VoIP UA* to make another call to the GSM phone number of the dual-mode device (indicated by “Step 4” and “Step 5”).

Once the *GSM Phone* module on the dual-mode device is alerted for the incoming call, the *Caller ID* is exposed to the *Handoff Control* module (“Step 6”). Based on its internal state, the *Handoff Control* module matches the caller to the remote peer that the WiFi mode is currently in communication with, and then directs the *GSM Phone* module to answer the call using the *Answer* command (“Step 7”). After the call between the remote peer and the GSM mode is established and the voice stream is flowing between two ends (“Step 8”), the remote peer can then send a *Handoff Complete* command back to the dual-mode device (“Step 9”). The *Handoff Control* module on the dual-mode device then responds with *Handoff Complete ACK* (“Step 10”) and directs the *VoIP UA* module to hang up the call (“Step 11”), thus completing the handoff process from WiFi to GSM modes at the dual-mode device. It can be observed that the handoff process is transparent to the user, and the two voice streams can overlap in time without temporal discontinuity. The prolonged call setup time between GSM and VoWLAN as shown in Section 5 thus can be properly handled without introducing undesirable gap.

We note that while the protocol handshakes shown in Fig. 3 require the remote peer to make the new call (to call the *GSM Phone* module on the dual-mode device), it is possible to let the dual-mode device make the new call instead. As shown in Fig. 4, after the *Handoff Control* module on the dual-mode device receives the acknowledgement to the *Handoff Trigger* command, it may direct the *GSM Phone* module to make a call to the remote peer (“Step 4” and “Step 5”). The remote peer can be made to answer the GSM call automatically (“Step 6” and “Step 7”) since the GSM phone number of the dual-mode device can be conveyed to it in the *Handoff Trigger* command to be maintained in the internal states of the *Handoff Control* module. The *Handoff Complete* command and acknowledgement, and the call release procedures follow as usual. The reason for designing such a variant of the handoff process is in consideration of the special billing structure of certain mobile telephony service providers. In many GSM billing plans both the caller and the callee share the cost of communication, while in many others *only the caller is billed*

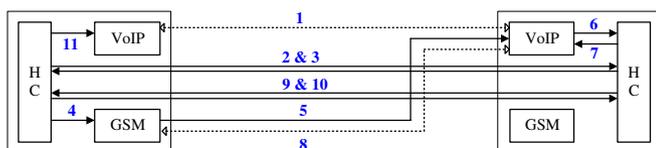


Fig. 4. Vertical handoff from WiFi to GSM (dual-mode device makes the second call).

for the call made. Therefore, it may matter in some scenarios whether the remote peer or the dual-mode device makes the new call. Depending on the billing plans and usage preference, during the *Handoff Trigger* handshakes of “Step 2” and “Step 3” in Figs. 3 and 4, the remote peer and the dual-mode device can negotiate who should make the new call for vertical handoff, and then continue with the call flows of the remaining handoff process accordingly.

4.3. Handoff from GSM to WiFi

While the dual-mode user may handoff from WiFi to GSM modes due to loss of WiFi service, it is also possible that the user handoffs from GSM to WiFi modes due to availability of the low-cost WiFi service. Such migration may be determined between two users on an end-to-end basis subject to user preference and the nature of the call. In this section, we present the handoff process when the dual-mode decides to migrate its GSM call to VoWLAN call when it has access to WiFi connectivity.

As shown in Fig. 5(a), the dual-mode device is initially in a call with a remote VoIP UA through the GSM mode (“Step 1”). Once the *WiFi Monitor* module notifies availability of WiFi service, and the *Handoff Control* module decides to initiate the handoff, the *Handoff Trigger* command and acknowledgement are exchanged between the two devices through IP service (“Step 2” and “Step 3”). The *Handoff Control* module on the dual-mode device conveys the identity of the VoIP user agent (the “number” that the *GSM Phone* module dials) on the remote peer so the *Handoff Control* module on the remote peer can associate the incoming message with the existing call. If it is desirable that the remote peer makes the new call, then the *Handoff Control* module on the remote peer directs the *VoIP UA* to make a call to the *VoIP UA* on the dual-mode device as shown in “Step 4” and “Step 5” of Fig. 5(a). The required information for dialing (e.g. SIP URI of the dual-mode device) is conveyed through the *Handoff Trigger* command sent earlier. The dual-mode device can then match the *Caller ID*, and answers the call (“Step 6” and “Step 7”). After the new call (between two VoIP user agents) is established, the old GSM call may be released as indicated in “Step 9,” “Step 10,” and “Step 11” in the figure.

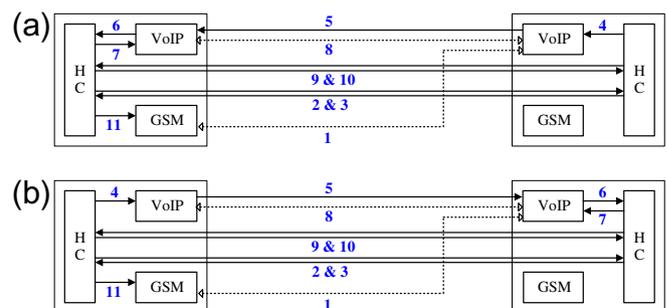


Fig. 5. Vertical handoff from GSM to WiFi. (a) Remote peer makes the second call. (b) Dual-mode device makes the second call.

If, on the other hand, it is desirable that the dual-mode device makes the new call, then the *Handoff Control* module on the dual-mode device directs its *VoIP UA* module to make the call to the remote SIP UA as shown in “Step 4” and “Step 5” of Fig. 5(b). The remote SIP UA matches the *Caller ID* to the states maintained in the *Handoff Control* module, and answers the call. The rest of the handoff process follows as usual.

In summary, the proposed vertical handoff procedure aims to ensure “make-before-break” soft handoff while reusing the VoIP UA and GSM phone modules on the dual-mode mobile device. Through handshakes between the two *Handoff Control* modules on the end devices, the setup of the new call and the release of the old call can be performed on an end-to-end basis without modifying

existing GSM and SIP infrastructure. We show in Fig. 6 the detailed protocol handshakes between the dual-mode device (D) and the remote SIP peer (R) for WiFi-to-GSM (Fig. 3) and GSM-to-WiFi (Fig. 5(a)) vertical handoffs. For clarity, the figure collocates the *Handoff Control* module and the *VoIP UA* module on the remote SIP client. It is clear from the figure that the two SIP proxies and the SIP gateway do not need to be changed for handling the handoff related messages between the two *Handoff Control* modules.

5. Performance evaluation

In this section, we present testbed results to show the performance of the proposed solution for vertical handoffs.

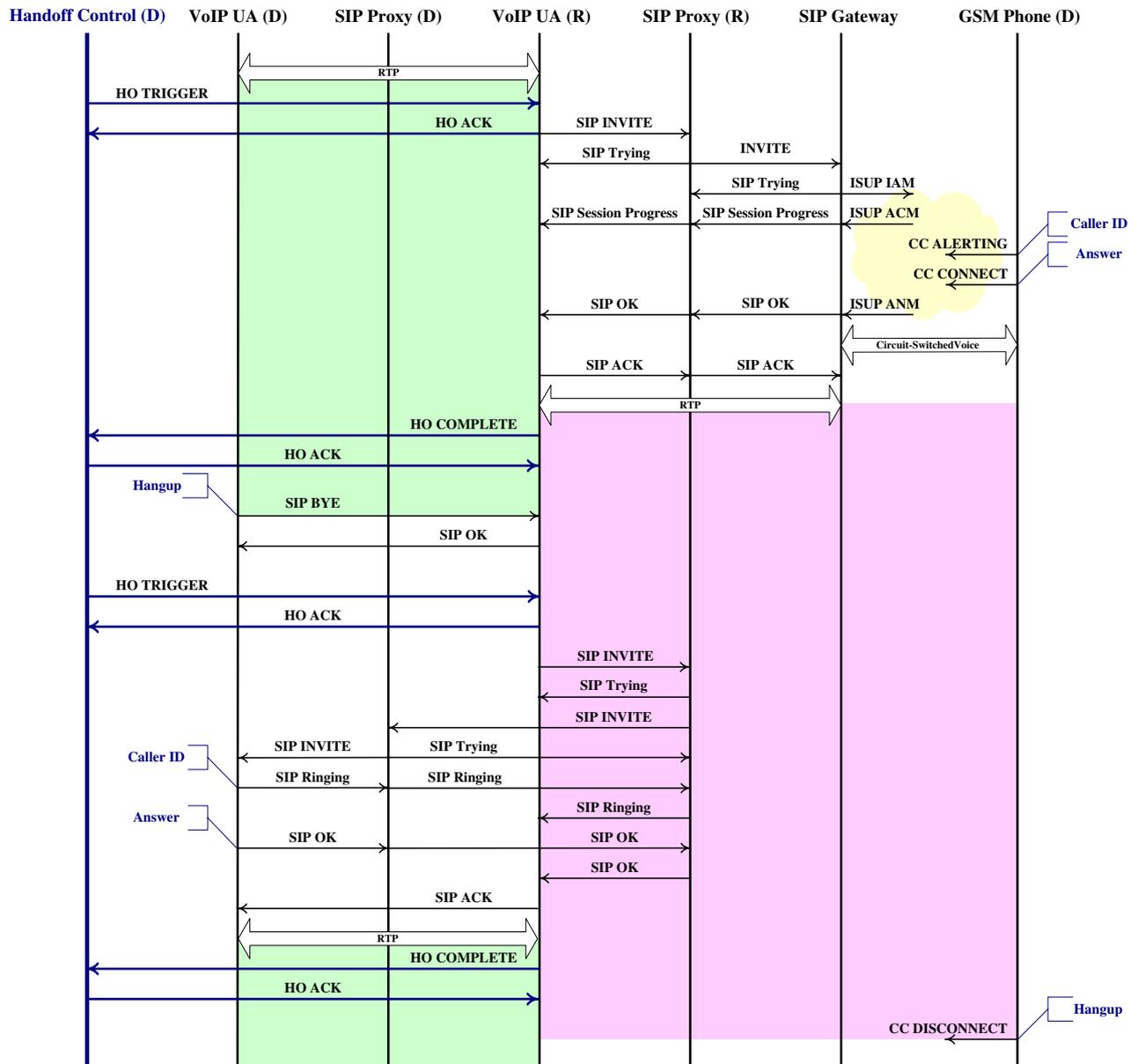


Fig. 6. Vertical handoff diagram between WiFi and GSM.

5.1. Testbed setup

We use a GSM PDA with a built-in 802.11b radio interface as a dual-mode mobile device. The PDA is equipped with the Intel PXA272 416 MHz CPU and runs the Windows Mobile 5.0 operating system. We port the open-source sipXtapi library [19] to the Windows Mobile 5.0 platform and use it to develop the SIP user agent for dialing and receiving VoWLAN calls on the PDA. As shown in Fig. 7, the SIP infrastructure in the testbed consists of an SER SIP proxy server and a Cisco SIP–PSTN gateway so phone calls between PSTN (and GSM) phones and SIP phones are possible. Another desktop computer (with wired connection) that runs the open-source sipXezPhone user agent [19] is used as the remote SIP phone. In the testbed, the dual-mode PDA has access to both city-wide GSM service and campus WiFi service, and hence it can make phone calls with the remote SIP phone using either mode of communication. During testbed experiments, the dual-mode PDA is carried in and outside the coverage of the WiFi service area to trigger vertical handoffs between WiFi and GSM modes.

In the following, we first use the testbed to measure the call setup time of GSM and VoWLAN phone calls and motivate the need for “make-before-break” soft handoffs in the target scenario. We then present the results during vertical handoffs from the WiFi to GSM modes.

5.2. Network characteristics

We use the testbed to measure the call setup time between the GSM/WiFi modes and the remote SIP client. The call setup time is measured at the caller from the time the call is placed to the time the voice from the callee is heard at the caller. It can represent the delay between the time the handoff is triggered until the time the voice stream of the new call arrives. We first measure the call setup time when the call is made from the WiFi mode on the dual-mode PDA, and then measure the time when the call is made from the GSM mode. Fig. 8 thus shows the call setup times of the two calls.

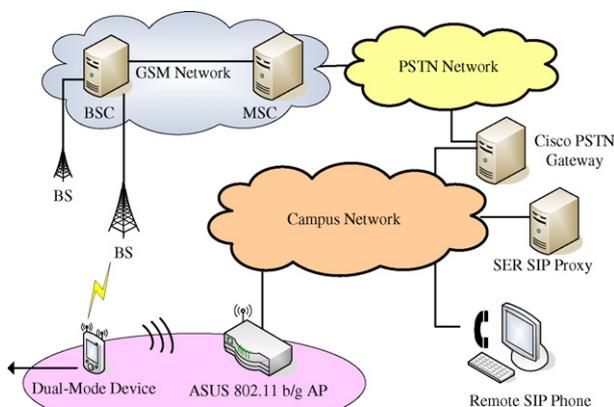


Fig. 7. Testbed setup.

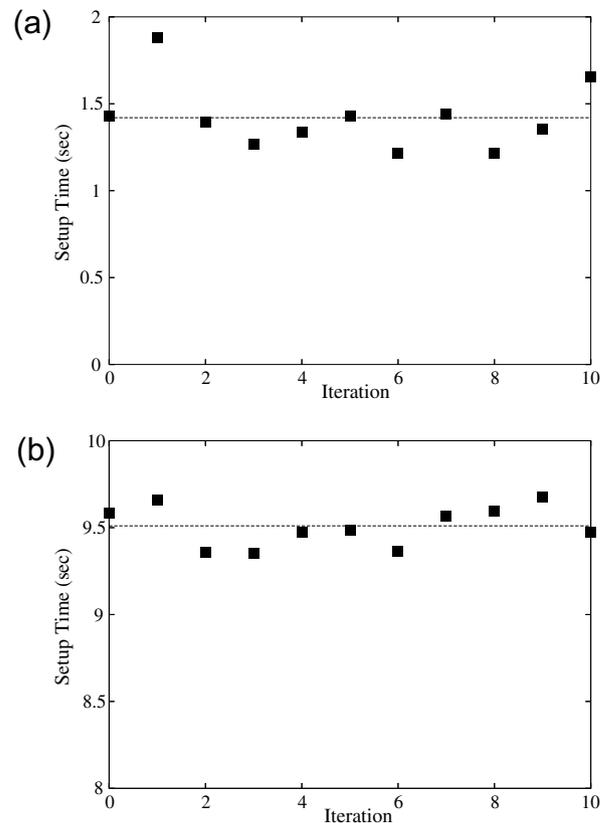


Fig. 8. Call setup time between the SIP phone and the dual-mode PDA. (a) VoWLAN call. (b) GSM call.

As shown in the figure, the call setup time for VoWLAN calls (made from the WiFi mode of the dual-mode device) ranges from 1.21 to 1.88 s, with an average of 1.42 s over 11 iterations. For GSM calls (made from the GSM mode of the dual-mode device), the call setup time ranges from 9.35 to 9.68 s with an average of 9.51 s. Since the call setup time is measured on the dual-mode device, it may involve delays due to the software and hardware processing time on the device. However, the primary reason for the prolonged delay between the GSM mode and the remote SIP phone is due to sophisticated signaling process (e.g. ISDN User Part signaling as shown in Fig. 6) and traversals of multiple networks involving the GSM access and core network, PSTN network, Internet, and the campus network.

To understand the call setup time when the calls are established between two dual-mode devices, we replace the remote SIP phone with another dual-mode PDA and measure the call setup delay. Fig. 9 shows the delay when one of the device is operated in the GSM mode, and the other device uses either the WiFi or GSM mode to make the call. For calls established between one WiFi mode and one GSM mode, the maximum setup delay measured is 7.12 s, the minimum delay is 6.78 s, and the average delay is 6.94 s. For calls established between two GSM modes of the two devices, the maximum is 10.43 s, the minimum is 9.04, and the average is 9.83 s. It can be observed that call setup times involving the GSM network are quite long, and

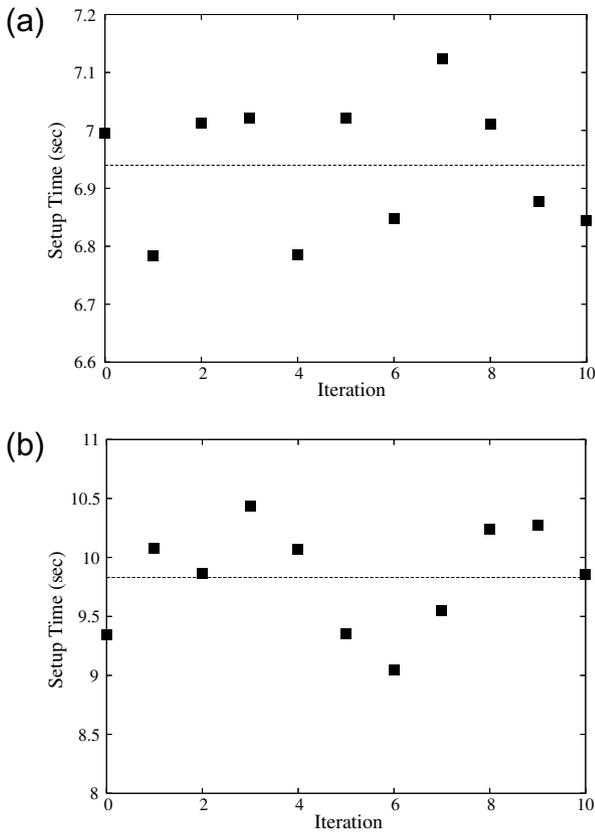


Fig. 9. Call setup time between two dual-mode PDAs. (a) WiFi to GSM. (b) GSM to GSM.

hence if conventional “break-before-make” or “consultation hold” call transfer is used for call migration on the dual-mode device, then the user will suffer from undesirable interruption. A “make-before-break” soft handoff manifests itself as a better approach for the target scenario.

5.3. Vertical handoff

We show the performance of the proposed solution when the dual-mode device migrates from the WiFi mode to the GSM mode. Since the GSM voice is not packetized, it is not possible to show the packet sequence number progression during handoffs. Instead, we show the speech waveform recorded on the dual-mode device during handoff.

In Fig. 10, we show the voice heard by the user of the dual-mode device when an active call is migrated from

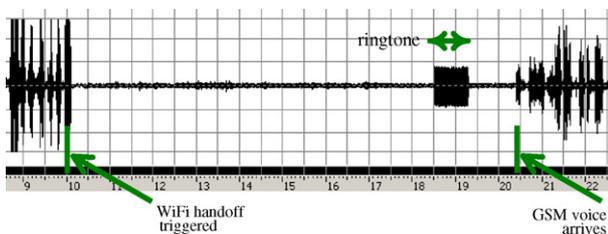


Fig. 10. Break-before-make vertical handoff.

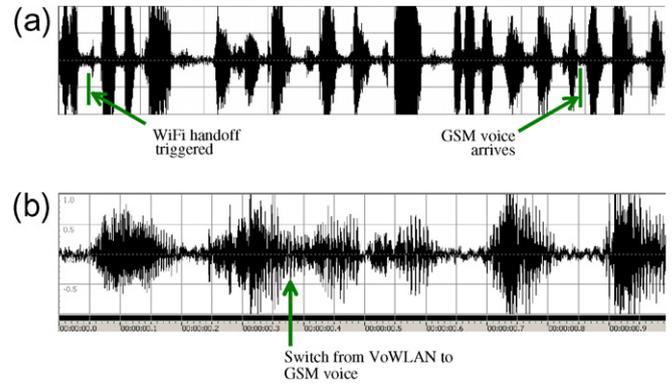


Fig. 11. Speech waveform during WiFi-to-GSM handoff. (a) Speech waveform. (b) Enlarged view during switching.

the WiFi mode to the GSM mode using “break-before-make” hard handoff. Specifically, after the handoff is triggered, the *Handoff Control* module on the dual-mode device negotiates with the remote peer for establishing the new GSM call. It, however, tears down the VoWLAN call as soon as the remote peer acknowledges the setup of the new call (blind transfer). As shown in the figure, due to the prolonged setup time of the GSM call, there is no overlap between the two voice streams, and the gap during the handoff is as long as 10 s. It is clear that such a gap would make the handoff unpleasant to the user of the dual-mode device.

Fig. 11(a), on the other hand, shows the voice heard at the dual-mode device using the handshakes as shown in Fig. 3. Since the GSM call is established while the VoWLAN call is still in session, the problem due to the prolonged setup time of the GSM call is minimized. The dual-mode device switches the audio path from the WiFi voice stream to the GSM voice stream after the GSM call is established, thus completing the handoff process. Fig. 11(b) shows an enlarged view of the time interval when the voice is switched from the WiFi mode to the GSM mode. It can be seen that no unnecessary audio gap is introduced when the switching takes place due to the “make-before-break” nature of the proposed approach.

6. Discussion

We have shown in earlier sections that it is possible to use end-to-end approaches for vertical handoffs across heterogeneous telephony systems. In this section, we discuss issues and future research directions.

6.1. Communication with the GSM peer

We have discussed in Sections 4.2 and 4.3 the vertical handoff procedures when the dual-mode device is communicating with another SIP soft phone (or the WiFi mode of another dual-mode device) and needs to migrate the call between the GSM and WiFi modes. While we have restricted the focus of this paper to the aforementioned sce-

narios, the proposed approach can be extended to the scenario when the remote peer involves the GSM mode of another dual-mode device. Assuming that the call was initially established between the WiFi mode on one device and the GSM mode on the other device. During the progress of the call, if the device with the active WiFi mode is moved beyond the WLAN service area, a handoff needs to be triggered from the WiFi mode to the GSM mode – similar to that discussed in Section 4.2. However, since the remote peer operates in the GSM mode, it is not guaranteed to have IP connectivity for the two *Handoff Control* modules to exchange *Handoff Trigger* messages and related information as discussed in Section 4.2. As mentioned before, for the handoff to proceed without user intervention, the remote GSM peer needs to associate an incoming GSM call with the existing VoWLAN call for automatic call answer. While the association of the two identities (GSM phone number and SIP URI for example) on any dual-mode device can be published and made available through a central, public look-up repository, we show an alternate approach that does not rely on such a network provision.

As shown in Fig. 12(a), when the *Handoff Control* module on one dual-mode device receives the instruction to make a phone call with the GSM mode of the remote dual-mode device, it first sends an SMS text message containing the required binding information (and possibly user preference as discussed in Section 4.2) for potential vertical handoff (“Step 0”). The *Handoff Control* module on the remote dual-mode device parses the incoming SMS message and stores it in the internal states. Afterward, the GSM call between the two devices is setup as usual (“Step 1”). If it is decided that the call needs to be migrated from the WiFi mode to the GSM mode on one device, then the *Handoff Control* module can direct its *GSM Phone* to make the new call. The *GSM Phone* on the remote device can associate the incoming call with the dual-mode device it is communicating with using the internal states in the *Handoff Control* module. The automatic answer of the new call and release of the old call can then proceed as shown in Fig. 12(a). Similarly, for vertical handoffs from the GSM mode to the WiFi mode, through pre-establishment of the device information between two dual-mode devices, the call migration can also

be made transparent to the user as shown in Fig. 12(b). Our future work includes implementation and experimentation with the proposed approach when the call involves such a remote GSM peer.

6.2. DSP support

In the performance results shown in Section 5, when the voice stream of the call from the new mode arrives, the audio path is switched from the old voice stream to the new voice stream. While such a scheme is straightforward and can achieve acceptable performance in many scenarios (as shown in Fig. 11), in some scenarios it might be necessary to further optimize the performance through application of digital signal processing (DSP) techniques. In particular, since the end-to-end delays of the old and new paths (for GSM and VoWLAN calls) may be different, simply by switching from one voice stream to another may introduce acoustic artifacts such as jump or repetition of speech if the delay mismatch is larger than that can be absorbed by the playout buffer. To avoid this problem, it is helpful to perform “time alignment” of the two voice streams for finding the relative *time offset* of the two streams. Metrics such as cross-correlation or average magnitude difference function (AMDF) may be used for time alignment while incurring low computation complexity. In this way, the appropriate “glue” point from the new voice stream can be identified when the audio path is switched from the old to the new stream.

Fig. 13 presents the extended functional blocks of the dual-mode device with an integrated *Audio Mixing* module.

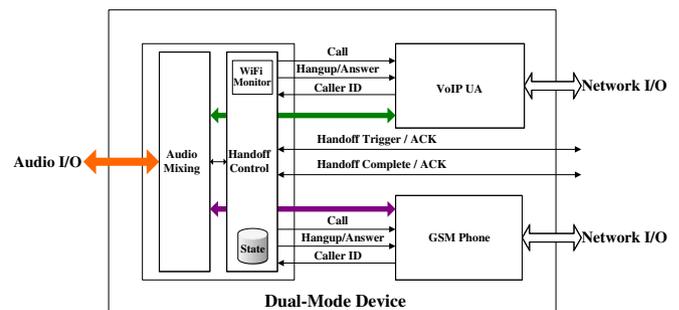


Fig. 13. Dual-mode device with the audio mixing module.

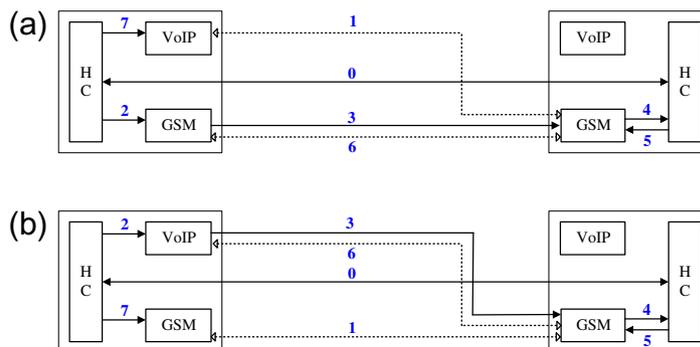


Fig. 12. Vertical handoff with a remote GSM peer. (a) WiFi to GSM. (b) GSM to WiFi.

The *Audio Mixing* module processes voice streams received from the *VoIP UA* and *GSM Phone* modules before they are sent to the system device for audio I/O. A more in-depth investigation of the optimization techniques for audio mixing is our ongoing work.

7. Conclusions

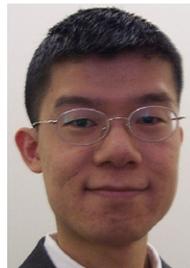
In this paper, we investigate the problem of supporting seamless voice communications across heterogeneous telephony systems. Instead of relying on interworking solutions, however, we consider a solution that can be deployed solely on the end devices. Specifically, we focus on VoIP service based on the Session Initiation Protocol (SIP), and we explain why existing solutions cannot be used. We identify the key challenges of such an end-to-end solution as the requirement to address the prolonged call setup time during vertical handoffs. Toward this goal, we propose an end-to-end approach for enabling “make-before-break” soft handoffs. We conduct testbed experiments using GSM–WiFi dual-mode PDA and investigate the voice quality when the call is migrated from VoWLAN to GSM. Evaluation results show encouraging results for future research along this direction.

Acknowledgment

This work was supported in part by funds from the National Science Council under Grant NSC-94-2213-E-002-052, and the Excellent Research Projects of National Taiwan University under Grant 95R0062-AE00-04.

References

- [1] Apple Inc., iPhone, <<http://www.apple.com/iphone/technology/specs.html>>, 2007.
- [2] Dopod International Corp., 838Pro, <<http://www.dopodasia.com>>, 2006.
- [3] Dopod, International Corp., CHT9100, <<http://www.dopod.com.tw>>, 2006.
- [4] M. Stemm, R. Katz, Vertical handoffs in wireless overlay networks, *ACM/Kluwer Mobile Networks and Applications* 3 (4) (1998) 335–350.
- [5] H.-Y. Hsieh, K.-H. Kim, Y. Zhu, R. Sivakumar, A Receiver-Centric Transport Protocol for Mobile Hosts with Heterogeneous Wireless Interfaces, *ACM MobiCom*, San Diego, CA, USA, 2003, pp. 1–15.
- [6] R. Inayat, R. Aibara, K. Nishimura, A seamless handoff for dual-interfaced mobile devices in hybrid wireless access networks, in: *IEEE International Conference on Advanced Information Networking and Applications (AINA)*, Fukuoka, Japan, Mar. 2004, pp. 373–378.
- [7] H.-H. Choi, O. Song, D.-H. Cho, A seamless handoff scheme for UMTS-WLAN interworking, in: *IEEE Globecom*, Dallas, TX, USA, 2004, pp. 1559–1564.
- [8] 3rd Generation Partnership Project, Generic access to the A/Gb interface; stage 2, 3GPP TS 43.318 V6.7.0, July 2006.
- [9] M. Buddhikot, G. Chandranmenon, S. Han, Y.-W. Lee, S. Miller, L. Salgarelli, Design and implementation of a WLAN/CDMA2000 interworking architecture, *IEEE Communications Magazine* 41 (11) (2003) 90–100.
- [10] A. Salkintzis, G. Dimitriadis, D. Skyrianoglou, N. Passas, N. Pavlidou, Seamless continuity of real-time video across UMTS and WLAN networks: challenges and performance evaluation, *IEEE Wireless Communications Magazine* 12 (3) (2005) 8–18.
- [11] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, SIP: Session Initiation Protocol, IETF RFC 3261 (2002).
- [12] A. Vemuri, J. Peterson, Session Initiation Protocol for Telephones (SIP-T): context and architectures, IETF RFC 3372 (2002).
- [13] H. Schulzrinne, E. Wedlund, Application-layer mobility using SIP, *ACM Mobile Computing and Communications Review* 4 (3) (2000) 47–57.
- [14] R. Sparks, The Session Initiation Protocol (SIP) refer method, IETF RFC 3515 (2003).
- [15] Motorola Inc., Motorola seamless mobility, <http://www.motorola.com/wlan/solution_cn620.html>.
- [16] S. Donovan, “The SIP INFO method, IETF RFC 2976 (2000).
- [17] R. Sparks, A. Johnston, D. Petrie, Session Initiation Protocol call control – transfer, IETF Internet Draft *draft-ietf-sipping-cc-transfer-07* (work in progress), 2006.
- [18] S. Quackenbush, T. Barnwell III, M. Clements, *Objective Measures of Speech Quality*, Prentice Hall, Englewood Cliffs, NJ, USA, 1988.
- [19] SIPfoundry, sipXtapi – user agent framework SDK, <<http://www.sipfoundry.org/sipxtapi.html>>.



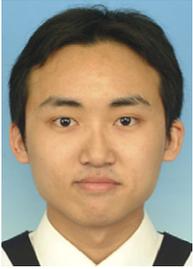
Hung-Yun Hsieh received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, ROC, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, Georgia, USA. He joined the Department of Electrical Engineering and the Graduate Institute of Communication Engineering at National Taiwan University as an Assistant Professor in August 2004. His research interests include wireless systems, mobile computing, and network protocols.



Chung-Wei Li received the B.S. degree in electrical engineering from National Central University, Jhongli, Taiwan, ROC, in 2004, and the M.S. degree in communication engineering from National Taiwan University, Taipei, Taiwan, ROC, in 2006. His research interests include voice over IP technology and wireless communication.



Shuo-Wei Liao received the B.S. degree in electrical engineering from National Central University, Jhongli, Taiwan, ROC. He is currently pursuing his M.S. degree in the Graduate Institute of Communication Engineering at National Taiwan University, Taipei, Taiwan, ROC. His research interests are in the area of video coding technologies and video streaming on wireless network as well as voice over IP networking technologies.



Yu-Wen Chen received the B.S. degree in electrical engineering from National Central University, Jhongli, Taiwan, ROC. He is currently pursuing his M.S. degree in the Graduate Institute of Communication Engineering at National Taiwan University, Taipei, Taiwan, ROC. His research interests include video compression, error resilience, and multimedia streaming.



Hsiao-Pu Lin received the B.S. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, ROC. He is currently pursuing his M.S. degree in the Graduate Institute of Communication Engineering at National Taiwan University, Taipei, Taiwan, ROC. His research interests are in the area of real-time multimedia streaming technologies over wired and wireless networks.



Tsung-Lin Tsai received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, ROC. He is currently pursuing his M.S. degree in the Graduate Institute of Communication Engineering at National Taiwan University, Taipei, Taiwan, ROC. His research interest is multimedia networking.