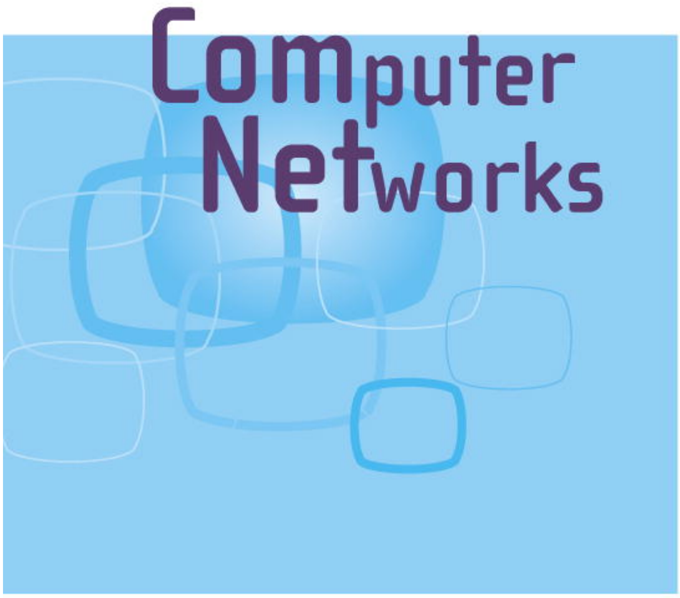
Volume 51 Issue 8 6 June 2007 ISSN 1389-1286



Available online at www.sciencedirect.com
ScienceDirect
Computer Networks 51 (2007) 2049–2066
Computer Networks
www.elsevier.com/locate/comnet

On the design of mobility management scheme for 802.16-based network environment ☆

Chun-Chuan Yang [a,*], Chun-Shian Tsai [b], Junn-Yen Hu [a,b], Tzu-Chien Chuang [a,b]

[a] Department of Computer Science and Information Engineering, National Chi Nan University, #1, University Road, PULI, Nantao 545, Taiwan, ROC
[b] Department of Computer Science and Information Engineering, Chung Chou Institute of Technology, Taiwan, ROC

Received 2 May 2006; received in revised form 4 October 2006; accepted 13 October 2006
Available online 9 November 2006

Responsible Editor: W. Kellerer

Abstract

In this paper, we investigate the characteristics of IEEE 802.16 and conclude that it is better to equip BS (base station) and SS (subscriber station) with Layer 3 functionality. Therefore, an 802.16 network can act as the backbone network of different subnets for better deployment. Based on the two IEEE Specifications, 802.16-2004 and 802.16e, we propose two kinds of paradigms of the 802.16 network technology for mobile networking. In the first paradigm, a novel concept called middle-domain mobility management in between macro- and micro-domain for 802.16-2004 is proposed. The management scheme of middle-domain is designed to accommodate different micro-mobility protocols in an 802.16-2004 network environment. Moreover, a mathematical analysis and simulation study are presented for performance evaluation. In the second paradigm, by comparing with traditional overlay networks (e.g. GPRS/WLAN), we have found that the characteristics for the 802.16e/802.11 overlay network are actually different from traditional overlay networks. To provide more efficient vertical handoff, a novel protocol called speed-based vertical handoff scheme (SVH) is proposed. A Simulation study has demonstrated that SVH can achieve a better performance than its WLAN-first counterpart in terms of less signaling and fewer packet losses.

Keywords: 802.16; BWA; Mobility management; Wireless overlay network; Vertical handoff

1. Introduction

Broadband Wireless Access (*BWA*) technology provides an easy, time-saving, and low-cost method for deployment of next generation (beyond 3G) network infrastructure. Since 1998, IEEE 802.16 working group has launched a standardization process called *Wireless Metropolitan Area Network* (*Wireless*

☆ This work was supported in part by the National Science Council, Taiwan, R.O.C., under grant no. NSC94-2219-E-260-004.
* Corresponding author. Tel.: +886 49 2910960x4887; fax: +886 49 2915226.
E-mail addresses: ccyang@csie.ncnu.edu.tw (C.-C. Yang), s2321903@ncnu.edu.tw (C.-S. Tsai).

1389-1286/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2006.10.007

MAN™) for BWA. The newly released specification of 802.16 (*IEEE Std 802.16-2004*) [1] focuses on fixed location wireless access and can support up to 134 Mbps bit rate. Moreover, IEEE 802.16 working group is currently working on the standardization of a new 802.16 interface, 802.16*e* [2], to support wireless access with high mobility. The *WiMax Forum* (*Worldwide Interoperability for Microwave Access*) [3], a wireless industry consortium with about 100 members including major vendors such as *AT&T*, *Fujitsu*, *Intel*, and *Siemens Mobile*, is supporting 802.16 technology and promoting its commercial use, which means 802.16 is becoming the most important technology in BWA.

As illustrated in Fig. 1, a typical 802.16 network consists of a *base station* (*BS*) and a couple of *subscriber stations* (*SS*) that connect to the BS via a high-speed wireless link. The BS acts as a gateway to the Internet. Legacy LANs or even more complex subnet systems can connect to the 802.16 network via SS. An 802.16 network (including the Legacy LANs that connect to SS) can cover a large geographical area since the distance between BS and SS can be up to 30 miles (in the case of 802.16-2004).

Similar to other 802 protocols, IEEE 802.16 defines the specification in physical layer (Layer 1) and MAC layer (Layer 1.5). Thus, from the viewpoint of layering architecture in networking, an 802.16 network is basically a subnet and the BS or SS acts as a Layer 2 (L2) device (bridge, for instance). However, it is improper to view an 802.16 network as a subnet like 802.3 or 802.11 LAN, since (1) an 802.16 network can cover a large geographical area and (2) a large number of users (including mobile hosts) in the network would cause serious performance degradation if the whole 802.16 network is only a single broadcast domain.

For example, *Address Resolution Protocol* (*ARP*) requires the ARP request frame to be broadcast in the whole 802.16 subnet in order to get the mapping from the logical IP address to the physical address.

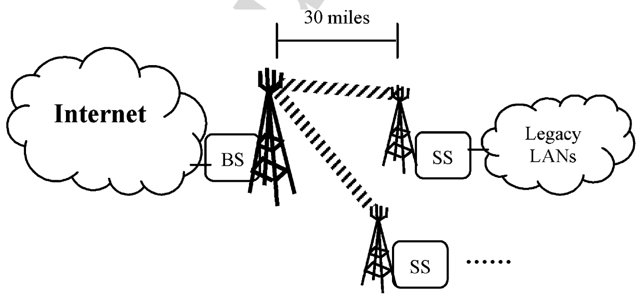


Fig. 1. General view of 802.16 network.

Moreover, in order to support mobile computing in 802.16, L2 mobility management as well as L2 handoff control [4] require the handoff frames to be broadcast in the network, creating more annoyed broadcast frames in the 802.16 network.

Therefore, we conclude that it is better to equip BS and SS with Layer 3 (L3) functionality such that 802.16 network acts as the backbone network of different subnets to enhance 802.16-based network deployment. This kind of network deployment (heterogeneous subnets interconnected by L3 802.16 BS/SS) is actually a form of internet, and it is called *802.16 network environment* in this paper.

There are two approaches to support mobility for users in an 802.16 network environment: (1) mobile hosts equipped with 802.11 interface roaming among WLANs or cellular systems that connect to SS (in the case of 802.16-2004), or (2) mobile hosts equipped with 802.16e interface connecting to the BS directly. In this paper, we aim to design appropriate mobility management schemes for each of the two approaches respectively as briefly explained in the following.

Deployment of 802.16 technology by approach (1) for mobility supporting is called the paradigm of "*802.16-2004 mobile network environment*" in this paper. Given that 802.16-2004 BS and SS are equipped with L3 functionality as discussed above, an 802.16-2004 mobile network environment is beyond the ability of a Layer 2 mobility management scheme. Hence, we investigated the feasibility of applying existing L3 mobility management schemes in the paradigm. We have found that the current two-tier mobility management (*macro-mobility* + *micro-mobility*) [5,6] cannot fit in 802.16-2004 mobile network environment well. Therefore, a new concept of *middle-domain* mobility management is proposed.

On the other hand, mobile hosts with 802.16e interface in approach (2) act like mobile subscriber stations that connect to 802.16e BS directly. In this case, there is only one single 802.16e cell for mobile hosts and it does not require elaborate mobility management. Thus, we are more interested in the extension of approach (2) in which WLANs are connecting to stationary SS and mobile hosts are equipped with both 802.16e and 802.11 interfaces. Deployment of the extension is called the paradigm of "*802.16e/802.11 overlay network environment*" in the paper. As in traditional overlay networks (e.g. GPRS/802.11 overlay networks) [7], mobility management for 802.16e/802.11 overlay network environment needs to deal with not only horizontal handoff (802.11 ⟨-⟩ 802.11) but also vertical handoff (802.16e ⟨-⟩ 802.11). Since the mobility management scheme proposed for 802.16-2004 mobile network environment can be applied for horizontal handoff, design of the vertical handoff scheme is the main focus for 802.16e/802.11 overlay network environment.

The rest of the paper is structured as follows. First of all, we make a brief survey of (1) existing L3 mobility protocols, (2) related work of mobility management for 802.16, and (3) vertical handoff schemes in overlay networks in Section 2. The concept of 802.16 middle-domain and the associated handoff mechanism for 802.16-2004 mobile network environment are presented in Section 3. In Section 4, an efficient vertical handoff scheme for 802.16e/802.11 overlay network environment is proposed. Finally, Section 5 concludes this paper.

2. Related work

2.1. L3 mobility management

Currently L3 mobility management solutions can be broadly classified into two categories: *macro-mobility* and *micro-mobility* management solutions, in which the movement of mobile users between two network domains is referred to as macro-mobility and the movement between two subnets within one domain is referred to as micro-mobility. In the following, we make a brief survey on the most typical macro-mobility protocol, *Mobile IP* (*MIP*) [8,9], and two typical micro-mobility protocols [10], *Cellular IP* (*CIP*) [11,12] and *Mobile IP Regional Registration* (*MIP-RR*) [13]. Moreover, specific issues about MIP such as *Route Optimization MIP* (*ROMIP*) [14], *MIPv6* [15,16] and *Fast MIPv6* (*FMIPv6*) [17,18] are also surveyed.

2.1.1. Macro-mobility protocol: MIP

In MIP, a *mobile host (MH)* uses two IP addresses: a fixed *home address* and a *care-of-address (CoA)* that changes at each new point of attachment (subnet). A router called *Home Agent (HA)* on an MH's home network is responsible for maintaining the mapping (binding) of the home address to the CoA. When an MH moves to a foreign network, the MH obtains a CoA from the *Foreign Agent (FA)* and registers the CoA with its HA. In this way, whenever an MH is not attached to its home network, the HA gets all packets destined for the MH and arranges to deliver to the MH's current point of attachment by *tunneling* the packets to the MH's CoA.

2.1.2. Micro-mobility protocols: CIP and MIP-RR

CIP is proposed to provide local mobility and handoff support for frequently moving hosts. It supports fast handoff and paging in CIP access networks. For mobility between different CIP networks, it can interwork with MIP to provide wide-area mobility support. A Cellular IP network consists of a gateway (GW) and base stations (BS). The gateway connects the Cellular IP network to Internet. Cellular IP base stations are nodes that have an interface to a wireless network and interfaces to the wired network. Packets transmitted from mobile hosts are always routed from the base station to the gateway by a hop-by-hop shortest path routing. On the other hand, packets destined to an MH reach the GW first. Then the GW forwards the packets to the MH using the host-specific routing path.

MIP-RR aims to reduce the number of signaling messages to the home network and also reduce the signaling delay by performing registrations locally in a regional network. When an MH first arrives at a regional network, it performs a home registration with its HA. During the home registration, the HA registers the CoA of the MH, which is actually a publicly routable address of another mobility agent called a *gateway foreign agent (GFA)*. When an MH changes FAs within the same regional network, it performs only a regional registration to the GFA to update its local CoA. The packets for the MH are first intercepted by its HA, which tunnels those to the registered GFA. The GFA checks its visitor list and forwards the packets to the corresponding FA of the MH. The FA further relays the packets to the MH. In order to enhance the efficiency of mobility management in MIP-RR, more levels of *Hierarchical Foreign Agent (HFA)* can be added between GFA and FA.

2.1.3. Specific issues: ROMIP, MIPv6 and FMIPv6

To remedy the problem of triangular routing and reduce the packet loss during handoff, ROMIP was proposed. ROMIP allows every CN to cache and use binding copies. The original binding for an MH is kept in its HA, but ROMIP supports that a binding copy can be propagated to the requiring nodes. Local bindings in a CN enable most packets in a traffic session to be delivered by direct routing.

Moreover, an MH also informs its previous FA about the new CoA, so that the packets tunneled to the old location (due to an out-of-date binding copy) can be forwarded to the current location. This forwarding mechanism in ROMIP reduces the handoff latency and thus reduces the packet loss during handoff. However, the improvement of ROMIP over MIP in terms of routing efficiency and smaller handoff latency is at the cost of significantly larger signaling overhead.

MIPv6 takes advantage of the larger address space of IPv6 as well as the idea of ROMIP for routing improvement to mobile hosts. Therefore, the CN supporting MIPv6 must maintain the binding cache for the communicating MH and perform binding update after MH handoff as in ROMIP. FMIPv6 was proposed to achieve the following two goals: (1) to allow the MH to send packets as soon as the MH detects a new subnet link and, (2) to deliver packets to the MH as soon as the new attachment is detected by the new access router. FMIPv6 achieves the goals by informing the MH of the new AR's advertised prefix and validating the prospective new CoA on the new link prior to MH movement. Furthermore, FMIPv6 sets up a bidirectional tunnel between the old AR and the MH at the new CoA to reduce the handoff latency, which is conceptually similar to the idea of the forwarding mechanism in ROMIP.

2.2. 802.16 mobility management

Recent trends about 802.16 mobility management have already aimed at the investigation on requirements of localized IP mobility management (IP-MM) [19]. Chow and Garcia [20] presented an integration of macro- and micro-mobility in Mobile IP for 802.16e network. A *domain access router* (*AR*) lies at the edge of the Internet. It separates the Internet from the subnetwork below. The subnetwork below the domain AR is considered as a single administrative domain. A domain may consist of a number of interconnected access routers including intermediate ARs and edge ARs. Connected to the edge ARs are access points (AP) or base stations (802.16e BS). An AP is considered a neighbor to a given AP if its geographic coverage area is adjacent to the given AP. Each AP within a domain is connected to its neighboring APs via wired interfaces. Packets destined for an MH are forwarded from the previous AP to the new AP through the direct wired link during an intra-domain handoff. Domain ARs are responsible for packet forwarding during the inter-domain handoff. However, this work cannot fit in a large-scale domain (with a large number of MHs) well, since (1) the whole domain is treated as a single (Layer 2) broadcast domain, and the binding update between APs generates a great amount of broadcast messages for intra-domain handoffs and (2) the use of direct wired link between neighboring APs implies a significant cost of deployment. By contrast, the proposed architecture of 802.16-2004 middle-domain is more appropriate for a large-scale network.

2.3. Vertical handoff schemes

A *wireless overlay network* structure [21] is defined as the combination of wireless network interfaces fitting into a hierarchy of overlapped networks, and the network interface with larger cell size is located at a higher level in the hierarchy. As displayed in Fig. 2, the *horizontal handoff* (shown in dash line) is the handoff between base stations using the same type of wireless network interface, and the *vertical handoff* (shown in solid line) is the handoff between different types of network technology. There are two kinds of vertical handoffs: the *upward vertical handoff* is a handoff to an upper-layer network in the hierarchy (usually with larger cell size and lower bandwidth), and the *downward vertical handoff* is a handoff to a lower-layer network that provides smaller cell size and higher bandwidth.

Recent trend of the research in wireless overlay networks suggests that WLAN and third-generation

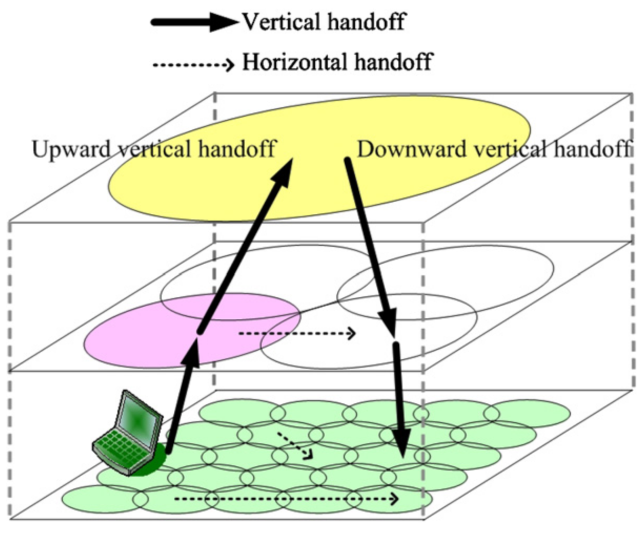


Fig. 2. The vertical and horizontal handoffs in wireless overlay network structure.

(3G) [22,23] such as CDMA2000 [24], GPRS [25], EDGE (Enhanced Data rates for GPRS Evolution) [26] and UMTS [27] co-exist to offer Internet access to end users. Ylianttila et al. [28] proposed a notion of *dwell-timer functionality* for vertical handoff decision between 802.11 and GPRS/EDGE using Mobile IPv6 to minimize the delay and maximize the throughput. In their scheme, the MH prefers using 802.11 interface and keeps monitoring the received signal level from the AP. The MH initiates the handoff to GPRS/EDGE while the received signal level during the dwell time is below the predefined threshold. Chang [29] proposed a mobility management scheme combining hierarchical micro-mobility with fast handoff and vertical handoff mechanisms to reduce the overheads associated with fast moving users. However, decision making for vertical handoff in Chang's work is merely based on received signal strength. As an extension of received signal strength-based vertical handoff, Nie et al. [30] also considered available bandwidth for vertical handoff.

Other than received signal strength-based mechanisms, there are a variety of researches for the vertical handoff scheme aiming at different goals in the literature. A cost function combining a variety of user- and network-valued metrics was define by Zhu and McNair [31]. They concluded that the network that results in the lowest value of the cost function is the network that would provide the most benefit to the users. Sharma et al. [32] proposed a vertical handoff scheme that are based on the monitored signal strength, quality, and noise levels of the wireless network interface. Two threshold levels of average signal strength, *low watermark* and *high watermark*, were adopted to avoid oscillating handoffs between GPRS and WLAN interfaces. Chen et al. [33] proposed an adaptive vertical handoff scheme associated with system discovery and handoff decision mechanisms. The proposed system discovery method was aimed to balance the power consumption and system discovery time. In the work of Wang et al. [34], a *policy-enabled* handoff system was proposed. The primary goal of the work is to balance the bandwidth load across networks with comparable performance. They claimed that a stabilized policy-based handoff system can achieve load balancing and improve network performance.

In this paper, as will be presented in Section 4, the difference between traditional overlay networks such as GPRS/WLAN and 802.16e/802.11 overlay network is investigated. In light of the different characteristics, a novel vertical handoff scheme considering the impact of the behavior of mobile hosts on signaling overhead is proposed.

3. Mobility management in 802.16-2004 mobile network environment

3.1. Problems of fitting existing protocols in 802.16-2004

Currently the two-tier mobility management uses the macro-mobility protocol and micro-mobility protocols at the same time but in different levels. The operation range of the macro-mobility protocol (MIP) is called the macro-domain and the operation range of a micro-mobility protocol such as CIP or MIP-RR is called a micro-domain in this paper.

It is proposed that 802.16-2004 devices are equipped with L3 functionality, and from the viewpoint of mobility management, 802.16-2004 devices are used to connect different micro-domains to Internet, thus 802.16-2004 is something between macro-domain and micro-domain as displayed in Fig. 3. There are two straightforward ways to design mobility management in an 802.16-2004 network environment: (1) macro-domain coupling, or (2) micro-domain coupling, as explained in the following.

3.1.1. Macro-domain coupling

As illustrated in Fig. 4, we can simply treat 802.16-2004 devices as part of the macro-domain. In such case, BS as well as SS of 802.16-2004 can get rid of mobility management and act just like regular routers. Moreover, in addition to functions of micro-mobility protocol, the *gateway router (GR)* of each micro-domain is equipped with MIP FA functions and is responsible for MIP home

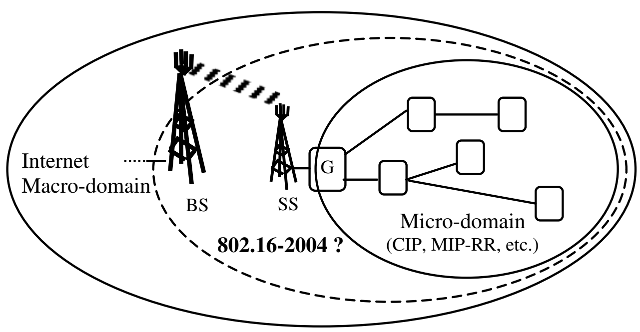


Fig. 3. Position of 802.16-2004 in mobility management.

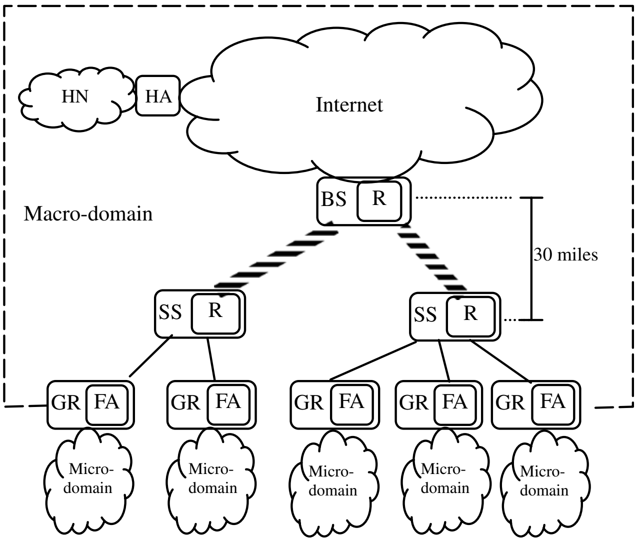


Fig. 4. Coupling 802.16-2004 with the macro-domain.

registration. However, this kind of coupling introduces performance problem in mobility management, since MIP home registration is required for the handoffs between different micro-domains in the same 802.16-2004 environment, which is inappropriate from the viewpoint of efficiency.

3.1.2. Micro-domain coupling

We can also treat the whole 802.16-2004 network environment as a single micro-domain, in which CIP or MIP-RR can be applied to support mobility management. Fig. 5 shows the typical examples for applying CIP and MIP-RR in 802.16-2004 respectively. As illustrated in Fig. 5(a), since CIP requires all data packets to be routed to the gateway (BS of 802.16-2004 in the case) before being routed to the destination, it results in a bad consequence that for internal traffic of which the source MH and the destination MH of the data packets are in the same 802.16-2004 network, the traffic is routed to the BS first, even if the two mobile hosts are on neighboring subnets. The performance problem is called "*lengthy internal data path*", which also results in the waste of precious link bandwidth between BS and SS.

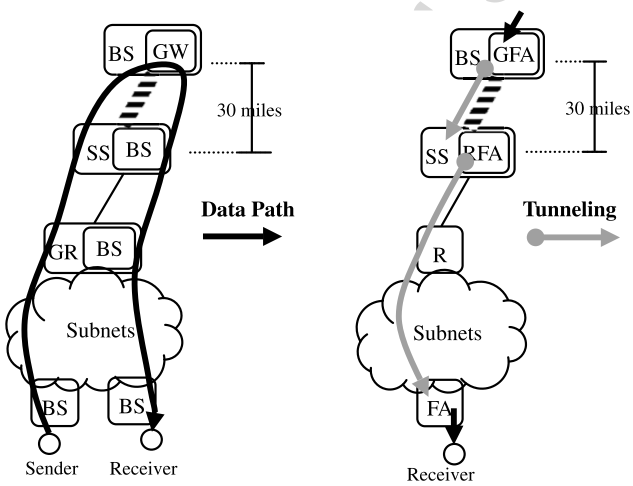
Fig. 5(b) shows the case of applying MIP-RR in an 802.16-2004 network environment, in which each SS is equipped with the function of RFA to avoid the problem of lengthy internal data path. However, this case introduces another type of deployment problem called "*one-hop tunneling*" between BS and SS. Since the idea of tunneling in mobility support is used for packet transmission across networks (regular routers) that are not supporting mobility scheme, one-hop tunneling is inappropriate and inefficient.

Last but not least, since an 802.16-2004 network environment can cover a large area, treating it as a single micro-domain lacks for the flexibility of adopting different mobility protocols in different micro-domains.

3.2. 802.16-2004 middle-domain mobility management

3.2.1. Basic idea

We conclude from the discussion in Section 3.1 that it is not appropriate to treat an 802.16-2004 network environment as part of the macro-domain nor a single micro-domain. Therefore, the idea of *middle-domain* emerges. Introducing the 802.16-2004 middle-domain results in a three-tier mobility management as illustrated in Fig. 6. Given that



(a) Lengthy internal data path in CIP (b) One-hop tunneling in MIP-RR

Fig. 5. Coupling 802.16-2004 with a micro-domain.

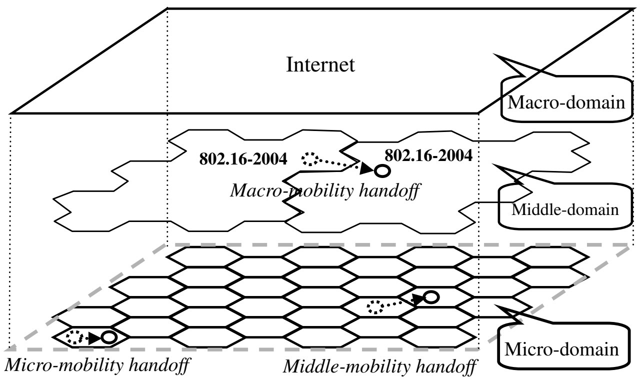


Fig. 6. The idea of 802.16-2004 middle-domain.

the idea of middle-domain is created after the two-tier mobility management, the operations of 802.16-2004 middle-domain are designed to be transparent from the viewpoint of macro- or micro-domains. That is, neither MIP nor micro-mobility protocols is required to be aware of the existence of the middle-domain, and the operations of MIP as well as micro-mobility protocols remain the same.

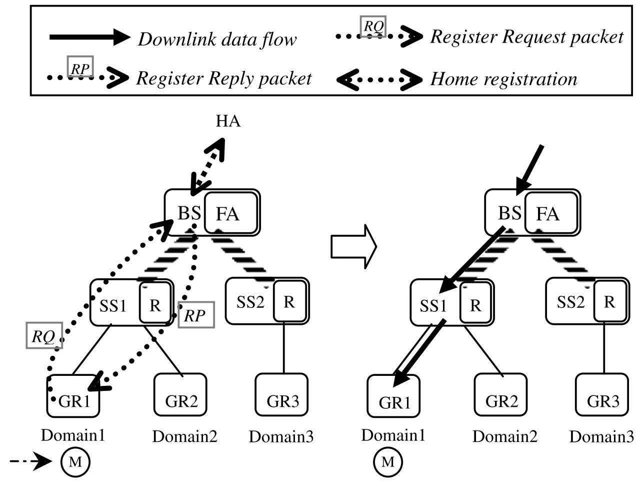
3.2.2. Location management

As in the two-tier mobility management, the micro-domain gateway router (GR) under each SS is required to equip with MIP FA functions and is responsible for MIP home registration. But the registration requests issued by the GR are intercepted by SS or BS in order to perform proper actions of the middle-domain. If an MH enters the middle-domain the first time, the following actions are taken in the middle-domain based on the intercepted MIP registration request:

1. The BS and SS en route create the location cache for the corresponding MH.
2. The BS allocates a *middle-domain CoA* (denoted by *M-CoA* in the paper) for the MH. The M-CoA is usually the address of the BS and is used in MIP registration.
3. The BS issues an MIP registration request with the M-CoA to the MH's HA on behalf of the GR. Meanwhile, the BS sends an MIP reply message back to the GR on behalf of the HA.

Signaling flow and data delivery for an MH entering the middle-domain the first time are illustrated in Fig. 7. In order to support middle-domain operations, the cache structures in BS and SS for an MH are displayed respectively in Fig. 8, in which MH's ID is the home address of the MH, the next hop for an MH in BS is the address of the next SS, the next hop for an MH in SS is the address of the next GR, the M-CoA is used in MIP home registration, and the micro-domain CoA is used in MIP reply to the GR.

From the viewpoint of the middle-domain, there are two types of handoff an MH can make after entering the middle-domain: (1) inter-micro-domain but intra-SS and (2) inter-SS but intra-BS. For case (1), the MIP registration request is intercepted by the SS. After updating the location cache for the MH, the SS sends an MIP reply message back to the GR as illustrated in Fig. 9(a). For case (2), the



(a) Signaling flow (b) Data delivery

Fig. 7. Mobile host enters 802.16-2004 middle-domain the first time.

M-CoA: Middle-domain CoA

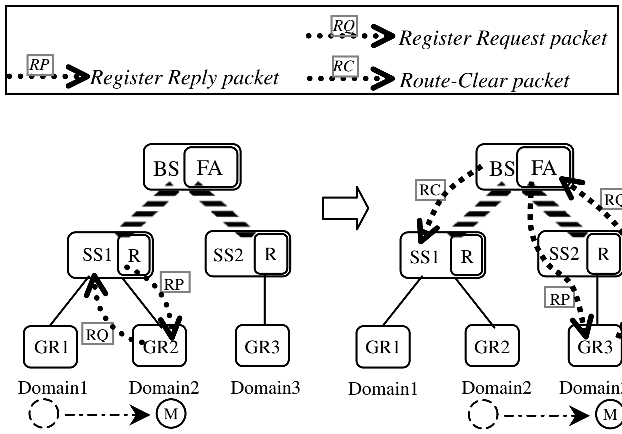
BS:

MH's ID	*Next Hop*	*M-CoA*	*Micro-domain CoA*

SS:

MH's ID	*Next Hop*	*Micro-domain CoA*

Fig. 8. Cache structure in the middle-domain.



(a) Intra-SS handoff (b) Inter-SS, Intra-BS

Fig. 9. Handoff scheme in 802.16 middle-domain.

BS and SS en route update/create the location cache respectively. The BS intercepts the MIP registration request and sends an MIP reply message back to the GR. Moreover, in order to forward internal data packets correctly, the location cache for the MH in the previous SS must be cleared. Therefore, a new control packet called *route-clear packet* is defined in the middle-domain. Signaling flow for inter-SS but intra-BS handoff is illustrated in Fig. 9(b). Note that there is no need to perform MIP home registration for both cases.

3.2.3. Data delivery

Data delivery from the CN to an MH with the introduction of the middle-domain is explained as follows. As illustrated in Fig. 10, data packets destined to an MH's home address are first intercepted by the HA. Since the CoA registered for the MH is the M-CoA, the HA tunnels the packets to the BS that allocated the M-CoA. The BS decapsulates the received packets and forwards them to the correct GR according to the location cache maintained by the BS and SS. Lastly, forwarding of the packets within a micro-domain is based on the operations of the micro-mobility protocol, which can be either tunneling-based (e.g. MIP-RR) or routing-based (e.g. CIP).

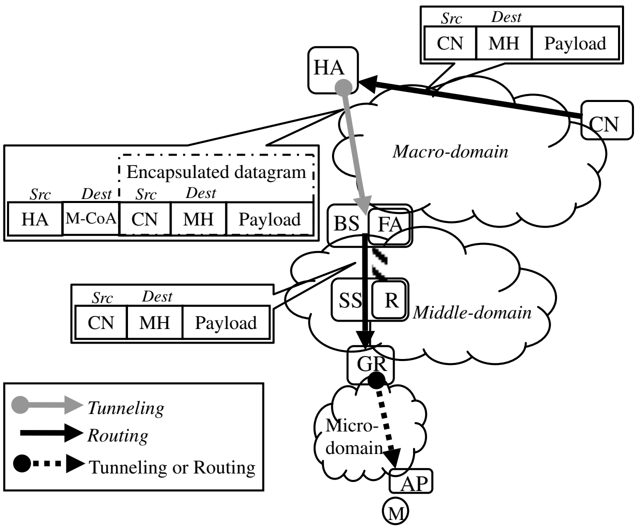


Fig. 10. Data delivery in three-tier mobility management.

Similar to CIP, data packets transmitted by an MH in 802.16-2004 are forwarded towards BS. However, the handling of the internal data flow is more efficient in the middle-domain as explained in the following. Since BS and SS maintain the location cache for each mobile host and if the data packets are destined to another MH in the same 802.16-2004, the crossover BS/SS of the source micro-domain and the destination micro-domain will identify the corresponding location cache for the destination MH and relay the data packets to the correct next hop.

3.3. Simulation study

Two 802.16-2004 networks are created in the simulation. As illustrated in Fig. 11, there is only one BS in each network. Four subscriber stations are connected to the BS, and four micro-domains are connected to each SS. There are 3200 mobile hosts in the network. In the beginning of the simulation, there are 100 mobile hosts located in each micro-domain. Time is slotted in the simulation and each mobile host leaves its current micro-domain and moves to one of the neighboring micro-domains

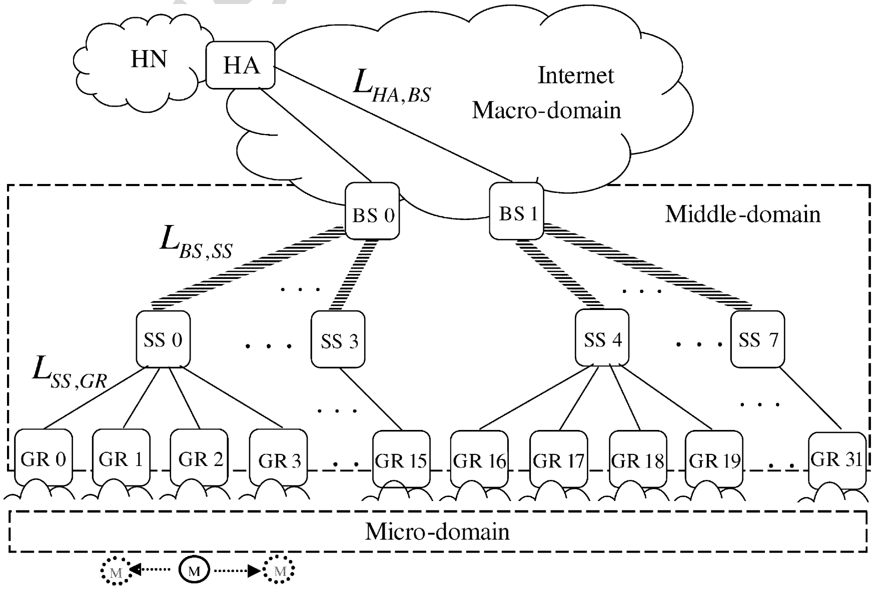


Fig. 11. Simulation environment.

with probability 20% for every time slot. Three handoff types are identified in the simulation. The total run time is 1000 time slots. Details of the simulation parameters are displayed in Table 1.

Two performance criteria are defined for comparing the proposed Middle-domain scheme and the MIP-applied contrast (i.e. the case without middle-domain): (1) the *average signaling cost* and (2) the *average handoff latency*. The average signaling cost is the registration cost for a handoff, and the average handoff latency is defined as the time to complete binding update after a handoff.

Fig. 12 shows the average signaling cost for three handoff types under different schemes. The average signaling cost including all handoff types is displayed in Fig. 13. The figures have demonstrated that the Middle-domain scheme can significantly reduce the signaling cost over the contrast (53% is saved in Fig. 13).

Performance of the average handoff latency is displayed in Figs. 14 and 15. Since the MH is more likely to perform "type 3" handoff (i.e. Intra-SS handoff), the average latency per handoff of the Middle-domain scheme is much shorter than that of the contrast.

Table 1
Simulation parameters

MH# =3200,100 per micro-domain 1000 time slots handoff probobility = 0.2 horizontal move	
Signal cost unit	$S_{\text{HA-BS}}$, $S_{\text{BS-SS}}$, $S_{\text{SS,GR}}$ (1 cost unit each)
Delay time	$L_{\text{HA-BS}} = 100$ Time Unit
	$L_{\text{BS-SS}} = 3$ Time Unit
	$L_{\text{SS-GR}} = 2$ Time Unit
Handoff types	Type 1: Inter BS
	Type 2: Inter SS, Intra BS
	Type 3: Inter GR, Intra SS

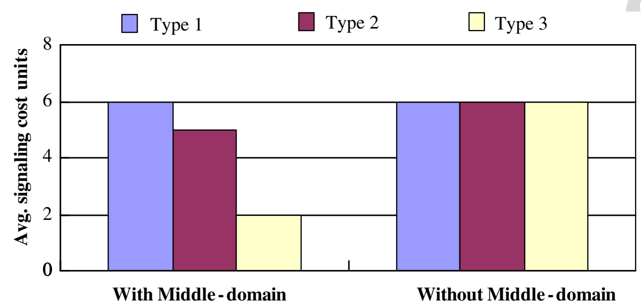


Fig. 12. Average signaling cost for three handoff types.

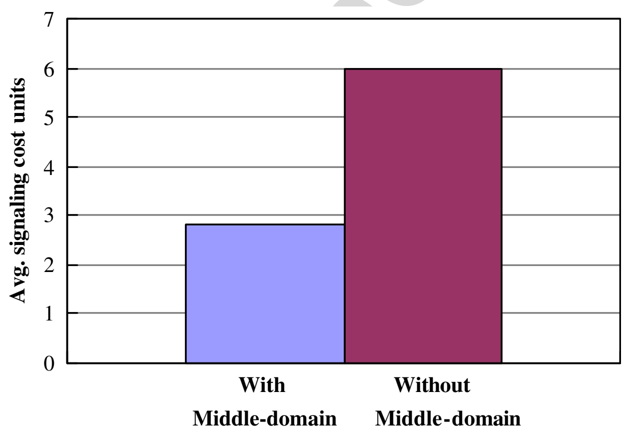


Fig. 13. The average signaling cost.

3.4. Theoretical analysis

In this section, a mathematical analysis of the signaling cost and the handoff latency for three different schemes, the Middle-domain scheme, the MIP-applied scheme, and the CIP-applied scheme, is presented. The theoretical environment for the analysis is displayed in Fig. 16. We assume there are x BSs located in the core network (Internet), each BS comprises y SSs, and each SS comprises z GRs (or FA). As in Fig. 17, we assume that the

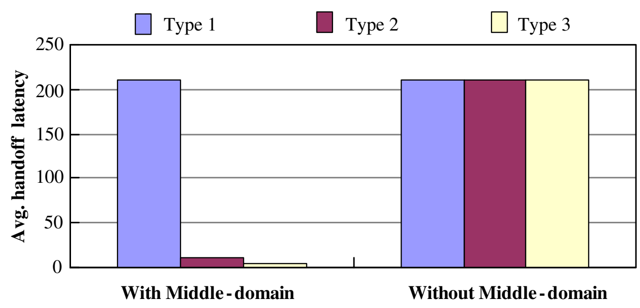


Fig. 14. Average handoff latency for three handoff types.

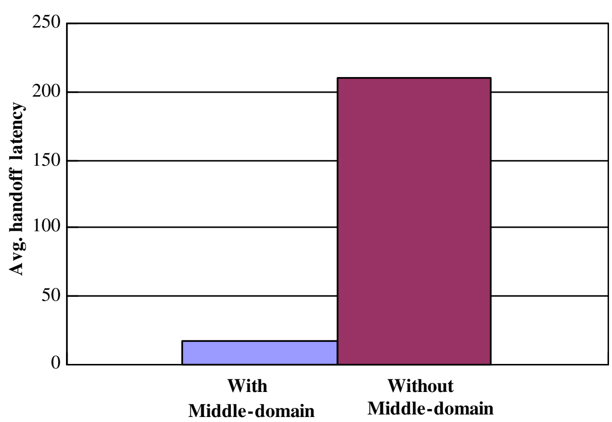


Fig. 15. The average handoff latency.

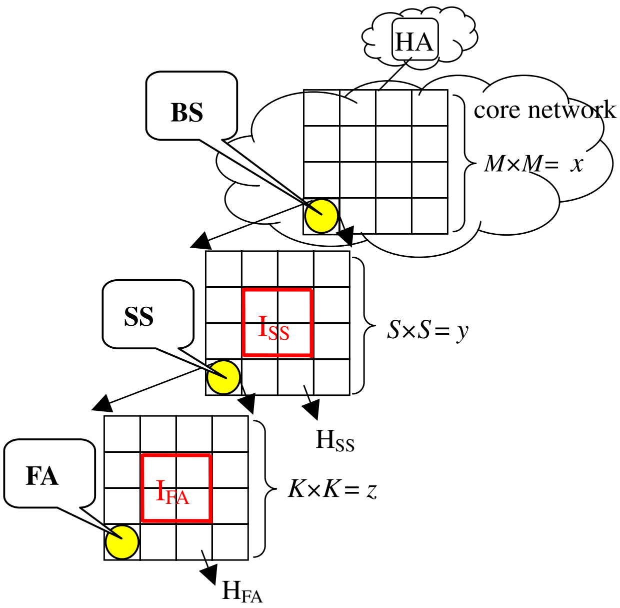


Fig. 16. Theoretical environment.

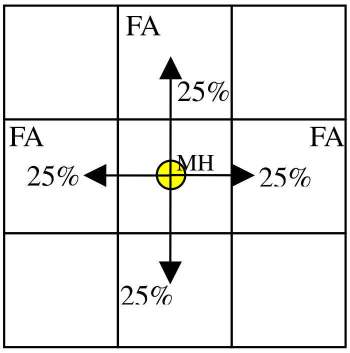


Fig. 17. Equivalent probability for an MH handoff.

probability for an MH hands over to one of four neighbor FAs is equivalent.

3.4.1. Probability of different handoff types

Consider the case of FA handoff, the probability of moving out of the SS area for each FA crossing is displayed in Fig. 18. The FAs in a single SS comprise inner FAs and peripheral FAs. The number of inner FAs in a single SS is denoted by I_{FA}, and the number of peripheral FAs is denoted by H_{FA}. We can easily derive the value of I_{FA} and H_{FA} as follows:

$$\begin{cases} I_{FA} = (K-2)^2 = (\sqrt{z}-2)^2, \\ H_{FA} = K^2 - I_{FA} = 4(\sqrt{z}-1). \end{cases} \quad (1)$$

Similar to formula (1), the number of inner SSs (I_{SS}) and peripheral SSs (H_{SS}) under a single BS can also be easily derived as follows:

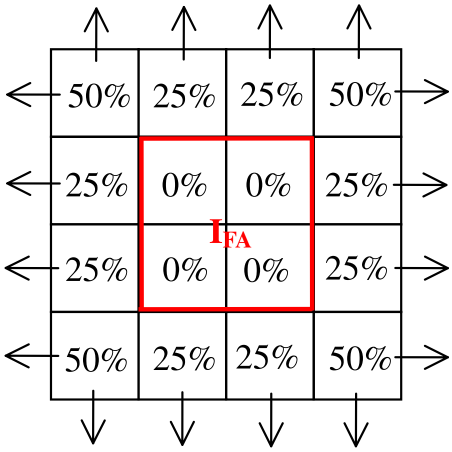


Fig. 18. Probabilities of moving out of the SS area for each FA crossing.

$$\begin{cases} I_{SS} = (S-2)^2 = (\sqrt{y}-2)^2, \\ H_{SS} = S^2 - I_{SS} = 4(\sqrt{y}-1). \end{cases} \quad (2)$$

As illustrated in Fig. 18, an FA handoff in each of the four corner FAs will move the MH out of the current SS with 50% probability. The probability of moving the MH out of the current SS for FA handoff in non-corner peripheral FAs is 25%.

Thus, the probability of an FA handoff making the MH stay in the same SS (denote by $P_{\text{Intra-SS}}$) is calculated as follows:

$$\begin{aligned} P_{\text{Intra-SS}} &= \text{Prob.[An FA handoff results in Non-SS crossing]} \\ &= \text{Prob.[Non-SS crossing]} \\ &= [4 \times 50\% + (H_{FA} - 4) \times (1 - 25\%) \\ &\quad + I_{FA} \times 100\%]/z = \frac{z - \sqrt{z}}{z}. \end{aligned} \quad (3)$$

Moreover,

$$\begin{aligned} &\text{Prob.[An FA handoff results in SS crossing]} \\ &\quad = \text{Prob.[SS crossing]} \\ &\quad = 1 - \text{Prob.[Non-SS crossing]} \\ &\quad = 1 - (z - \sqrt{z})/z = \sqrt{z}/z. \end{aligned} \quad (4)$$

The probability ($P_{\text{Inter-SS,Intra-BS}}$) of an FA handoff moving the MH out of the SS but still in the same BS is calculated as follows:

$$\begin{aligned} P_{\text{Inter-SS,Intra-BS}} &= \text{Prob.}[\text{An FA handoff results in SS crossing but Non-BS crossing}] \\ &= \text{Prob.}[\text{SS crossing} \cap \text{Non-BS crossing}] \\ &= \text{Prob.}[\text{Non-BS crossing} \mid \text{SS crossing}] * \text{Prob.}[\text{SS crossing}] \\ &= \{[4 \times 50\% + (H_{SS} - 4) \times (1 - 25\%) + I_{SS} \times 100\%]/y\} \times \frac{\sqrt{z}}{z} \\ &= \left[(y - \sqrt{y})/y\right] \times \frac{\sqrt{z}}{z} = \left(y - \sqrt{y}\right)/\left(y\sqrt{z}\right). \end{aligned} \quad (5)$$

Lastly, the probability ($P_{\text{Inter-BS}}$) of an FA handoff moving the MH out of the BS is calculated as follows:

$$\begin{aligned} P_{\text{Inter-BS}} &= \text{Prob.}[\text{An FA handoff results in BS crossing}] \\ &= \text{Prob.}[\text{BS crossing}] \\ &= \text{Prob.}[\text{BS crossing} \mid \text{SS crossing}] * \text{Prob.}[\text{SS crossing}] \\ &= \{1 - \text{Prob.}[\text{Non-BS crossing} \mid \text{SS crossing}] * \text{Prob.}[\text{SS crossing}] \\ &= \{1 - \left[(y - \sqrt{y})/y\right]\} * \left[\sqrt{z}/z\right] \\ &= \left(\sqrt{y}\sqrt{z}\right)/(yz). \end{aligned} \quad (6)$$

We could easily verify that $P_{\text{Intra-SS}} + P_{\text{Inter-SS,Intra-BS}} + P_{\text{Inter-BS}} = 1$, which demonstrating the correctness of the equations.

3.4.2. Signaling cost and handoff latency

For the MIP-applied scheme, the registration request is issued from MH to its HA and the HA replies the request to MH. Thus, the average signaling cost per handoff in the MIP-applied scheme is calculated as follows:

$$\begin{cases} S_{\text{MIP}} = 2d_{\text{HA,FA}}(R_{\text{core}}W_{\text{core}} + R_{\text{BS,SS}}W_{\text{BS,SS}} + R_{\text{SS,FA}}W_{\text{SS,FA}}), \\ 1 = R_{\text{core}} + R_{\text{BS,SS}} + R_{\text{SS,FA}}, \end{cases} \quad (7)$$

where S_{MIP} is the average signaling cost per handoff in the MIP-applied scheme (number of hops). $d_{\text{HA,FA}}$ is the average distance between home agent and foreign agent in terms of the number of hops. R_{core} is the ratio of the number of hops in the core network to the total number of hops between a home agent and a foreign agent. $R_{\text{BS,SS}}$ is the ratio of the number of hops between BS and SS to the total number of hops between a home agent and a foreign agent. $R_{\text{SS,FA}}$ is the ratio of the number of hops between SS and FA to the total number of hops between a home agent and a foreign agent. W_{core} is the weight of each hop in the IP core network. $W_{\text{BS,SS}}$ is the weight of each hop in between BS and SS for local 802.16 access network. $W_{\text{SS,FA}}$ is the weight of each hop in between SS and FA for local 802.16 access network.

The average handoff latency in the MIP-applied scheme is calculated as follows:

$$L_{\text{MIP}} = 2(L_{\text{HA,BS}} + L_{\text{BS,SS}} + L_{\text{SS,FA}}), \quad (8)$$

where L_{MIP} is the average handoff latency in the MIP-applied scheme. $L_{\text{HA,BS}}$ is the average delay time of link in between HA to BS for a handoff. $L_{\text{BS,SS}}$ is the average delay time of link in between BS to SS for a handoff. $L_{\text{SS,FA}}$ is the average delay time of link in between SS to FA for a handoff.

Different types of FA handoff in the Middle-domain scheme result in different registration costs. Therefore, the average signaling cost per handoff in the Middle-domain scheme is calculated by summing up the products of the probability and the registration cost under each case. That is,

$$S_{\text{Middle}} = C_{\text{Intra-SS}}P_{\text{Intra-SS}} + C_{\text{Inter-SS,Intra-BS}}P_{\text{Inter-SS,Intra-BS}} + C_{\text{Inter-BS}}P_{\text{Inter-BS}}, \quad (9)$$

$$\begin{cases} C_{\text{Intra-SS}} = d_{\text{HA,FA}}(2R_{\text{SS,FA}}W_{\text{SS,FA}}), \\ C_{\text{Inter-SS,Intra-BS}} = d_{\text{HA,FA}}[2(R_{\text{BS,SS}}W_{\text{BS,SS}} + R_{\text{SS,FA}}W_{\text{SS,FA}}) + R_{\text{BS,SS}}W_{\text{BS,SS}}], \\ C_{\text{Inter-BS}} = 2d_{\text{HA,FA}}(R_{\text{core}}W_{\text{core}} + R_{\text{BS,SS}}W_{\text{BS,SS}} + R_{\text{SS,FA}}W_{\text{SS,FA}}). \end{cases}$$

The average handoff latency in the Middle-domain scheme is calculated as follows:

$$L_{\text{Middle}} = 2L_{\text{SS,FA}}P_{\text{Intra-SS}} + 2(L_{\text{BS,SS}} + L_{\text{SS,FA}})P_{\text{Inter-SS,Intra-BS}} + 2(L_{\text{HA,BS}} + L_{\text{BS,SS}} + L_{\text{SS,FA}})P_{\text{Inter-BS}}, \quad (10)$$

where S_{Middle} is the average signaling cost per handoff in the Middle-domain scheme (number of hops). $C_{\text{Intra-SS}}$ is the registration cost for MHs crossing FA areas under a single SS. $C_{\text{Inter-SS,Intra-BS}}$ is the registration cost for MHs crossing SS areas under a single BS. $C_{\text{Inter-BS}}$ is the registration cost for MHs crossing between BS areas.

In the CIP-applied scheme, the signaling cost for an FA handoff in the same BS is a route update packet sent to the gateway (BS). For an inter-BS handoff, the BS performs the 2-way home registration on behalf of the MH. Therefore, the average signaling cost per handoff in the CIP-applied scheme is calculated as follows:

$$S_{\text{CIP}} = C_{\text{Intra-SS}}P_{\text{Intra-SS}} + C_{\text{Inter-SS,Intra-BS}}P_{\text{Inter-SS,Intra-BS}} + C_{\text{Inter-BS}}P_{\text{Inter-BS}}, \quad (11)$$

$$\begin{cases} C_{\text{Intra-SS}} = C_{\text{Inter-SS,Intra-BS}} = d_{\text{HA,FA}}(R_{\text{BS,SS}}W_{\text{BS,SS}} + R_{\text{SS,FA}}W_{\text{SS,FA}}), \\ C_{\text{Inter-BS}} = d_{\text{HA,FA}}(2R_{\text{core}}W_{\text{core}} + R_{\text{BS,SS}}W_{\text{BS,SS}} + R_{\text{SS,FA}}W_{\text{SS,FA}}). \end{cases}$$

The average handoff latency in the CIP-applied scheme is calculated as follows:

$$L_{\text{CIP}} = (L_{\text{BS,SS}} + L_{\text{SS,FA}})P_{\text{Intra-SS}} + (L_{\text{BS,SS}} + L_{\text{SS,FA}})P_{\text{Inter-SS,Intra-BS}} + (2L_{\text{HA,BS}} + L_{\text{BS,SS}} + L_{\text{SS,FA}})P_{\text{Inter-BS}}, \quad (12)$$

where S_{CIP} is the average signaling cost per handoff in the CIP-applied scheme (number of hops). L_{CIP} is the average handoff latency in the CIP-applied scheme.

3.4.3. Numerical results

The average signaling cost and handoff latency per handoff under the parameter setting of Table 2 are displayed in Figs. 19 and 20 respectively. The figures bring out some performance characteristics of the three schemes as explained in the following:

1. When the cell size (802.16 size) is extremely small ($y * z = 1$), the operation of the Middle-domain scheme degenerates to the MIP-applied scheme. Thus, the signaling cost and handoff latency in the Middle-domain scheme are the same as in the MIP-applied scheme.
2. As the cell size increases, the Middle-domain scheme outperforms the MIP-applied scheme in terms of the average signaling cost as well as the handoff latency because of the effect of localized registration.
3. The CIP-applied scheme outperforms the Middle-domain scheme in terms of the average signaling cost since the greater effect of localized registration in the CIP-applied scheme. However, the average handoff latency in the CIP-applied is slightly longer than that of the Middle-domain scheme as the cell size is pretty large (e.g. $y * z = 1024$ and 4096). The reason is explained as follows. The probability of intra-SS ($P_{\text{Intra-SS}}$) handoff increases as the cell size getting large. The intra-SS handoff is finished as the route update packet reaching the SS in the Middle-domain scheme. On the other hand, the handoff is finished until the route update packet received by the gateway (BS) in the CIP-applied scheme, even in the case of intra-SS handoff, which on the average resulting in a longer handoff latency than that of the Middle-domain scheme.

Table 2
Parameter setting

$d_{\text{HA,FA}}$	16 hops	(y,z)	(1,1), (4,4), (4,16), (4,64), (4,256), (4,1024)
R_{core}	0.5	W_{core}	8
$R_{\text{BS,SS}}$	0.25	$W_{\text{BS,SS}}$	1
$R_{\text{SS,FA}}$	0.25	$W_{\text{SS,FA}}$	1
$L_{\text{HA,BS}}$	100 ms	$L_{\text{BS,SS}}$	3 ms
$L_{\text{SS,FA}}$	2 ms		

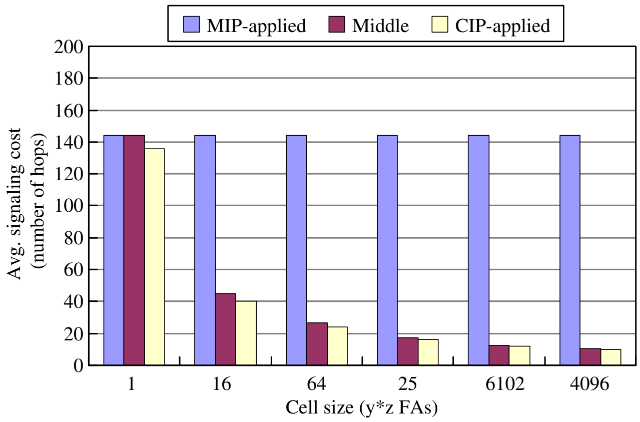


Fig. 19. The average signaling cost per handoff.

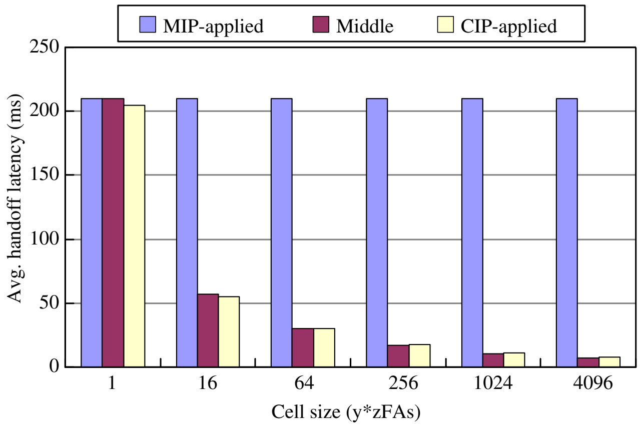


Fig. 20. The average handoff latency.

Table 3
The comparisons between 802.16e/802.11 overlay network and GPRS/WLAN overlay network

Network type	GPRS/WLAN	802.16e/802.11
Upper layer bandwidth	171.2 Kbps	15 Mbps
QoS supporting in upper layer	A few MHs	Many MHs
Lower layer bandwidth	802.11b 11 Mbps, 802.11g 54 Mbps	
Administrative business	Different	Same
Authentication process	A longer vertical handoff latency	A shorter vertical handoff latency
Mobility management	Heterogeneous	Homogeneous
Evaluate the cost for handoffs	Difficult	Easy

4. Mobility management in 802.16e/802.11 overlay network environment

4.1. Comparison with traditional overlay networks

An 802.16e/802.11 overlay network adopts two layers of overlapped network technologies, *WiMax* and *WiFi*, such that MHs equipped with both interfaces can have two different ways for network access. Unlike the traditional overlay network (e.g. GPRS/WLAN), an 802.16e/802.11 overlay network has different characteristics as listed in Table 3 and requires a new design of a more efficient vertical handoff scheme. The differences and comparisons between the 802.16e/802.11 overlay network and the GPRS/WLAN overlay network are presented in the following:

1. The upper layer of a traditional overlay network (GPRS) provides much lower capacity than the lower layer (WLAN) resulting that only a few mobile hosts can be supported at the same time via GPRS from the aspect of quality-of-service support. Therefore, a typical vertical handoff scheme in the GPRS/WLAN overlay network usually presents the preference for an MH to stay in the lower layer as long as possible and the upward vertical handoff (from WLAN to GPRS) is triggered only when the MH is out of the range of WLAN, which is called *WLAN-first vertical handoff scheme* in this paper. In contrast to GPRS, 802.16e can support much higher data rates up to 15Mbps, which means the idea of WLAN-first vertical handoff is no longer appropriate in the case of 802.16e/802.11 overlay networks and we should have different considerations in the design of the vertical handoff scheme.
2. In a traditional overlay network, different layers of network are often provided by different administrative businesses so that a longer vertical handoff latency is inevitable since the MH has to go through the authentication process during the vertical handoff. Moreover, it is difficult to evaluate and compare the cost of a vertical handoff with a horizontal handoff since different mobility management schemes are involved in different layers of network. By contrast, an integrated mobility management and security function for both layers can be provided in an 802.16e/802.11 overlay network since the same set of network devices (BS and SS) are in charge of both layers. Therefore, the vertical handoff in the 802.16e/802.11 overlay network is not as costly as in the GPRS/WLAN overlay network in terms of handoff latency and signaling overhead. Furthermore, the integrated mobility management also implies a common basis to evaluate the cost of vertical handoffs as well as horizontal handoffs in an 802.16e/802.11 overlay network.
3. The discussion above only addresses the characteristics of the network side. It is better to also consider the side of the mobile hosts in designing the vertical handoff scheme. Considering a high-speed MH moving around in an 802.16e/802.11 overlay network, it is better in term of handoff overhead to move the MH to the 802.16e layer instead of the 802.11 layer since the larger range of the 802.16e cell reduces the handoff frequency. By contrast, a low-speed MH should be more preferred to stay in the 802.11 layer than a high-speed MH. Therefore, the idea of *Speed-based Vertical Handoff scheme* (*SVH*) presented in the next section is proposed in this paper.

4.2. Speed-based vertical handoff scheme (SVH)

The main idea of SVH is that a high-speed MH always resides in the 802.16e layer to avoid the large number of handoff signaling packets, and a low-speed MH is preferred to reside (if possible) in the 802.11 layer to reduce competition over the 802.16e link capacity since traffic regulation can be exercised at the SS that in charge of 802.11 cells. Therefore, a threshold value of speed denoted by

V_1 is defined in SVH. Any MH with speed faster than V_1 is classified as high-speed and an upward vertical handoff is triggered to move the MH to the 802.16e layer. On the other hand, mobile hosts with speed less than V_1 reside in the 802.11 layer. In order to reduce handoff oscillation (so-called *Ping-Pong*effect) in which an MH with speed around V_1 moves back and forth in two layers in a very short time, another threshold value denoted by V_2 ($V_2 < V_1$) is defined for high-speed MHs to trigger downward vertical handoff. For a high-speed MH in the 802.16e layer, the downward vertical handoff is triggered when the speed of the MH is less than V_2. The algorithm of SVH is illustrated in Fig. 21.

In order to support SVH, the MH is required to have the ability of being aware of its moving speed. Given that *GPS* (*Global Positioning System*) is getting more cost-effective and popular in recent years, we assume that the high-end MH with both 802.16e and 802.11 interfaces is also equipped with the functionality of GPS to ease the task of speed estimation. Moreover, the routing cache is also required to associate with SVH for location management. Fortunately, mobility and location management for SVH can be built on top of the middle-domain mobility management scheme presented in Section 3.2 by merely adding a new type of cache entry at the BS for those MHs in the 802.16e layer. Mobility management for 802.11 MHs in the SVH scheme is apparently the same as in the middle-domain mobility management scheme, and the signaling packets used in the Middle-domain scheme are also used for SVH. We skip the detail of SVH operations to reduce the paper length.

4.3. Performance evaluation

4.3.1. Simulation environment and performance criteria

The network topology for the 802.16e/802.11 overlay network in the simulation is one BS and two SSs located in a 5000 m × 5000 m square. Each SS is in charge of 64 802.11 APs (Access Points) arranged in 8 × 8 mesh. The range of the 802.16e cell covers the whole simulation field. The range of each AP is a 200 m × 200 m square.

There are 300 mobile hosts in the network. In the beginning of the simulation, each MH is classified as one of the following types: *high-speed tendency*, *medium-speed tendency*, and *low-speed tendency*, and each type includes 100 MHs. Different probabilities for speed generation are associated with different types. As shown in Table 4, there are two intervals for speed generation: 0–49 km/h and 50–100 km/h. A high-speed tendency MH have the probability of 80% to generate its speed in between 50 and 100 km/h, and 20% for 0–49 km/h. Medium-speed tendency MHs have the same 50% probability for both intervals. For low-speed tendency MHs, 20% for 50–100 km/h and 80% for 0–49 km/h. *Random waypoint* algorithm with pause time = 0 is adopted for mobility model. Communication pairs

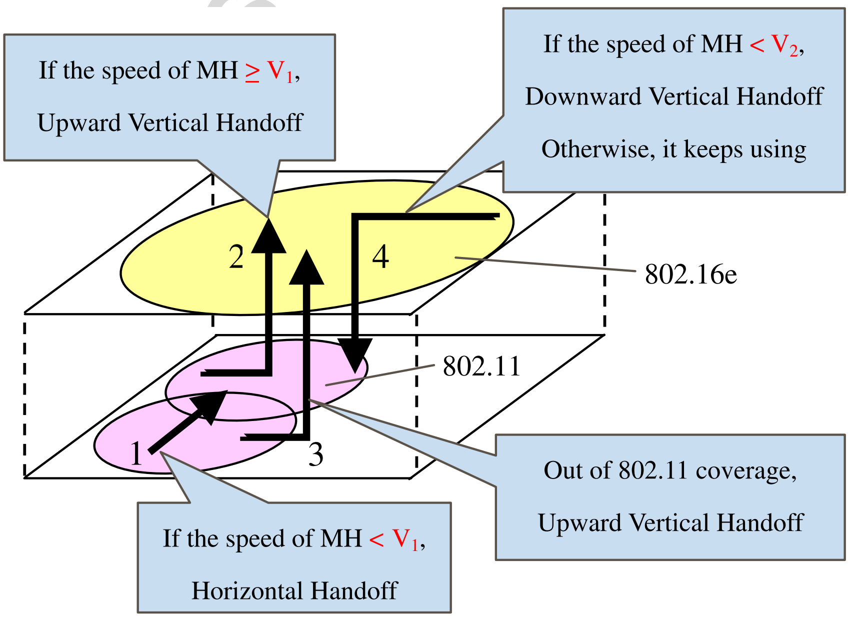


Fig. 21. Speed-based vertical handoff scheme.

Table 4
Speed selection for mobile host

Speed probability	Speed type		
	High-speed tendency (%)	Medium-speed tendency (%)	Low-speed tendency (%)
0–49 km/h	20	50	80
50–100 km/h	80	50	20

are randomly selected from the MHs in the simulation. The data rate for both directions in each communication pair is 10 packets/sec. The total simulation time is 3600 seconds.

Vertical handoff latency is set to be 500 ms and horizontal handoff latency is set to be 200 ms. The threshold values for V_1 and V_2 in SVH are set as 60 km/h and 40 km/h, respectively. Four performance criteria are defined for comparing SVH with WLAN-first vertical handoff scheme: (1) *average handoff times*, (2) *average handoff latency*, (3) *average packet loss*, and (4) *link usage*. The average handoff times are the total number of horizontal and vertical handoffs per MH. Average handoff latency per MH also includes both cases of vertical and horizontal handoffs, and the average packet loss is calculated as the average of the total number of lost packets for each MH during the simulation. Link usage for 802.16e or 802.11 is calculated as the total number of packets transmitted over the link per second.

4.4. Simulation results

Fig. 22 shows the average number of handoff for each type of mobile hosts and the summation of the average handoff times for SVH and WLAN-first vertical handoff scheme is displayed in Fig. 23 respectively. The figures have demonstrated that SVH can significantly reduce the number of handoff in comparing with the WLAN-first counterpart, since SVH moves high-speed MHs to the 802.16e layer to get rid of the large number of horizontal handoff in the 802.11 layer. More specifically, SVH saves up to 28%, 41%, and 77% of the handoff times over the counterpart for low-speed tendency, medium-speed tendency, and high-speed tendency MHs respectively. The higher reduction percentage for high-speed tendency MHs over the other two types is simply because a larger number of high-speed MHs are generated in the class. It is worth mentioning that the reduction of the handoff times also implies less signaling overhead by SVH.

Since the MH in SVH is more likely to perform vertical handoff, the average latency per handoff in SVH is longer than that of the WLAN-first vertical handoff scheme as displayed in Fig. 24. However, SVH reduces a large number of handoff as mentioned above, thus the average packet loss in SVH is much less than that of the WLAN-first scheme as displayed in Figs. 25 and 26. Lastly, as shown in Fig. 27, SVH introduces a moderate increase (8%) in the usage of the 802.16e link but reduces 21% of the usage of 802.11 link under the same

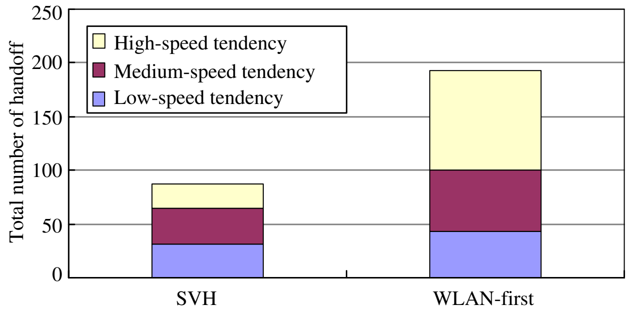


Fig. 23. Summation of the average handoff times.

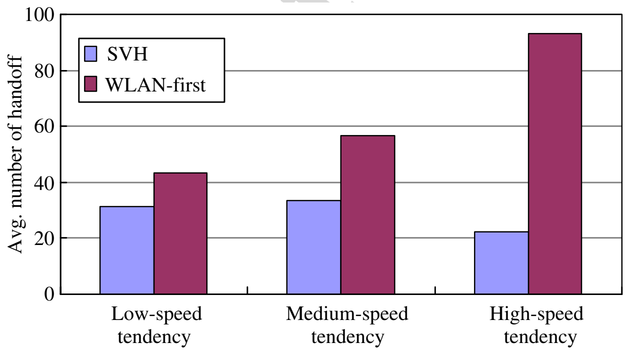


Fig. 22. Average number of handoff per mobile host.

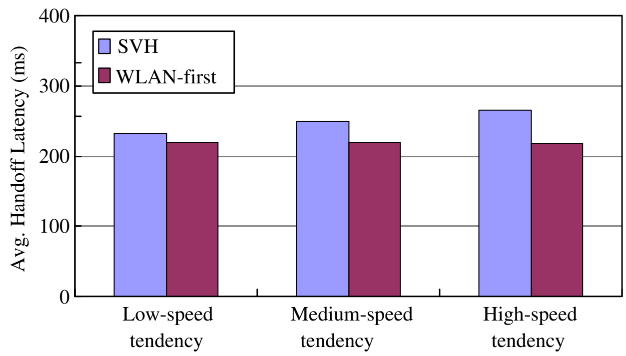


Fig. 24. Average handoff latency per handoff.

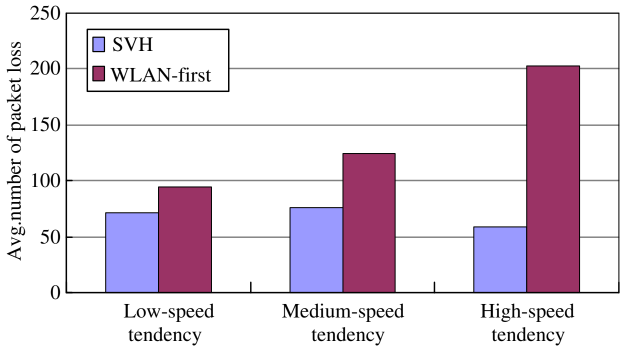


Fig. 25. Average number of packet loss per mobile host.

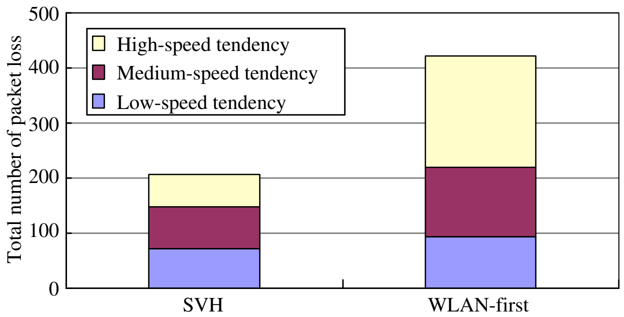


Fig. 26. Summation of the average packet loss of three MH types.

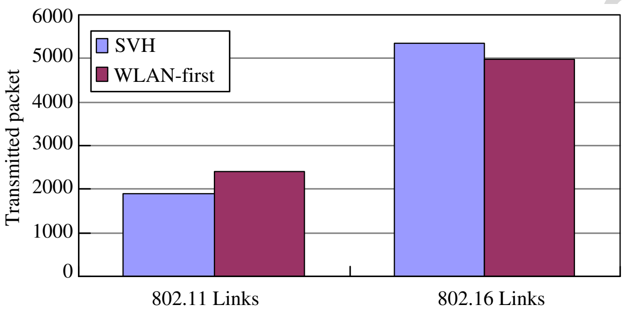


Fig. 27. Number of packets per second transmitted over the 802.16e/802.11 links.

input traffic, which demonstrates the improvement of transmission cost.

5. Conclusion

Broadband Wireless Access (*BWA*) technology provides an easy, time-saving, and low-cost method for deployment of next generation (beyond 3G) network infrastructure. The newly released specification of 802.16 (*IEEE Std 802.16-2004*) focuses on fixed location wireless access and can support up to 134 Mbps bit rate. Moreover, IEEE 802.16 working group is currently working on the standardization of a new 802.16 interface, 802.16*e*, to support wireless access with high mobility. Since an 802.16 network can cover a large geographical area and support a large number of users, we conclude that it is better to equip BS and SS with Layer 3 functionality such that 802.16 network acts as the backbone network of different subnets to enhance 802.16-based network deployment.

Based on the two specifications of IEEE 802.16, *802.16-2004* and *802.16e*, we propose two kinds of application for mobility supporting: (1) Paradigm of "*802.16-2004 mobile network environment*", in which the idea of *Middle-domain*mobility management as well as associated control mechanisms are proposed to support mobile users with 802.11 interface that connected to 802.16 subscriber stations and (2) Paradigm of "*802.16e/802.11 overlay network environment*", in which instead of the traditional WLAN-first vertical handoff, the *Speed-based Vertical Handoff* (*SVH*) scheme is proposed by considering the speed of mobile hosts to provide more efficient vertical handoff for mobile hosts with both 802.11 and 802.16e interfaces.

In summary, the contributions of the paper are listed as follows:

1. Deployment of 802.16 technology in the large scale has been discussed and Layer 3 deployment is recommended.
2. Two paradigms of 802.16 deployment for mobility supporting are proposed: *802.16-2004 mobile network environment* and *802.16e/802.11 overlay network environment*.
3. A novel concept namely *middle-domain mobility* management for 802.16-2004 mobile network environment is proposed. Middle-domain mobility management is designed to be able to accommodate different micro-mobility protocols and is transparent to macro-mobility and micro-mobility protocols. A mathematical analysis and simulation study show that middle-domain mobility management can achieve better performance in terms of less signaling cost and shorter handoff latency.
4. Differences between traditional overlay networks such as GPRS/WLAN and 802.16e/802.11 overlay network are presented. A novel protocol called *speed-based vertical handoff scheme (SVH)* is proposed by considering the impact of mobile host behavior on signaling overhead. Simulation study has demonstrated that SVH

outperforms WLAN-first vertical handoff scheme in terms of less control signaling and fewer packet loss.

References

[1] IEEE 802.16-2004, IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 1 Oct. 2004.

[2] IEEE 802.16 TGe, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, IEEE Standard No: 802.16E-2005 & 802.16/COR1, 28 Feb. 2006.

[3] S.J. Vaughan-Nichols, Achieving wireless broadband with WiMax, Computer (Jun.) (2004) 10–13.

[4] IEEE Std. 802.11f/D5, Draft Recommended Practice for Multi-Ventor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation, March 2003.

[5] M. Carli, A. Neri, A.R. Picci, Mobile IP and cellular IP integration for inter access networks handoff, in: Proceedings of IEEE International Conference on Communications (ICC), vol. 8, June 2001, pp. 2467–2471.

[6] I.F. Akyildiz, J. Xie, S. Mohanty, A survey of mobility management in next-generation all-IP-based wireless networks, IEEE Wireless Communications (Aug.) (2004) 16–28.

[7] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J.P. Makela, R. Pichna, J. Vallstron, Handoff in hybrid mobile data networks, IEEE Personal Communications 7 (2) (2000) 34–47.

[8] C.E. Perkins, Mobile IP, IEEE Communication Magazine 35 (5) (1997) 84–99.

[9] C.E. Perkins, IP Mobility Support for IPv4, IETF RFC 3344, August 2002.

[10] A.T. Campbell, J. Gomez, S. Kim, C.-Y. Wan, Comparison of IP micromobility protocols, IEEE Wireless Communications (Feb.) (2002) 72–82.

[11] A.T. Campbell, J. Gomez, An overview of cellular IP, in: Proceedings of IEEE wireless communications and networking conference (WCNC), vol. 2, Sept. 1999, pp. 606–610.

[12] A.T. Campbell, J. Gomez, S. Kim, A.G. Valko, C.Y. Wan, Z.R. Turanyi, Design, implementation, and evaluation of cellular IP, IEEE Personal Communications vol. 7 (issue 4) (2000) 42–49.

[13] Eva Gustafason, Annika Jonsson, C.E. Perkins, Mobile IPv4 regional registration, Internet draft, draft-ietf-mip4-reg-tunnel-01 (work in progress), November 2005.

[14] M. Dell'Abate, M. De Marco, V. Trecordi, Performance evaluation of mobile IP protocols in a wireless environment, in: Proceedings of IEEE International Conference on Communications (ICC), 1998, pp. 1810–1816.

[15] H.H. Choi, D.H. Cho, Mobility management based on mobile IP in mixed IPv4/IPv6 networks, in: Proceedings of 58th IEEE Vehicular Technology Conference (VTC), vol. 3, Oct. 2003, pp. 2048–2052.

[16] D. Johnson, C. Perkins, J. Arkko, Mobility Support in IPv6, RFC 3775, June 2004.

[17] R. Koodli, Fast Handovers for Mobile IPv6, RFC 4068, July 2005.

[18] L. Dimopoulou, G. Leoleis, I.S. Venieris, Introducing a hybrid fast and hierarchical MIPV6 scheme in a UMTS-IP converged architecture, in: Proceedings of 29th Annual IEEE International Conference on Local Computer Networks (LCN), Nov. 2004, pp. 42–49.

[19] Theo Pagtzis, Carl Williams, Charles Perkins, Peter Kirstein, Requirements for localised IP mobility management, in: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), vol. 3, 16–20 March 2003, pp. 1979-1986.

[20] Jocelyn Chow, Guillermo Garcia, Macro- and micro-mobility handoffs in mobile IP based MBWA networks, in: Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), vol. 6, 29 Nov.–3 Dec. 2004, pp. 3921–3925.

[21] M. Stemm, R.H. Katz, Vertical handoffs in wireless overlay networks, ACM Mobile Networking (MONET), Special Issue on Mobile Networking in the Internet, Summer, 1998, pp. 335–350.

[22] M. Buddhikot, G. Chandranmenon, S. Han, Y.W. Lee, S. Miller, L. Salgarelli, Integration of 802.11 and third-generation wireless data networks, in: Proceedings of the 22nd annual joint conference of the IEEE computer and communications societies (INFOCOM), vol. 1, 30 March, April 2003, pp. 503–512.

[23] Qian Zhang, Chuanxiong Guo, Zihua Guo, Wenwu Zhu, Efficient mobility management for vertical handoff between WWAN and WLAN, IEEE Communications Magazine 41 (11) (2003) 102–108.

[24] Wireless IP Network Standard. P.S0001-A-1, Third Generation Partnership Program 2 (3GPP2), 2000.

[25] Jian Cai, David J. Goodman, General packet radio service in GSM, IEEE Communications Magazine 35 (10) (1997) 122–131.

[26] A. Furuskar, M. Frodigh, H. Olofsson, J. Skold, System performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems, in: Proceedings of the 48th IEEE Vehicular Technology Conference (VTC), vol. 2, 18–21 May 1998, pp. 1284–1289.

[27] General Packet Radio Service (GPRS) Service Description (Stage 2), TS 122 060, ETSI, (2002).

[28] M.Ylianttila, M.Pande, J.Makela, P.Mahonen, Optimization scheme for mobile users performing vertical handoffs between IEEE 802.11 and GPRS/EDGE networks, in: Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), vol. 6, 25–29 Nov. 2001, pp. 3439–3443.

[29] Chung-Kuo Chang, A mobile-IP based mobility system for wireless metropolitan area networks, in: Proceedings of IEEE International Conference on Parallel Processing Workshops (ICPPW), June 2005, pp. 429–435.

[30] Jing Nie, Xin He, Zheng Zhou, ChengLin Zhao, Communication with bandwidth optimization in IEEE 802.16 and IEEE 802.11 hybrid networks, in: Proceedings of the 5th IEEE International Symposium on Communications and Information Technology (ISCIT), vol. 1, 12–14 Oct. 2005, pp. 27–30.

[31] Fang Zhu, Janise McNair, Optimization for vertical handoff decision algorithms, in: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), vol. 2, 21–25 March 2004, pp. 867–872.

[32] S. Sharma, I. Baek, Y. Dodia, Tzi-cker Chiueh, OmniCon: A mobile IP-based vertical handoff system for wireless LAN and GPRS links, in: Proceedings of 33rd International

Conference on Parallel Processing Workshops (ICPPW), 15–18 August 2004, pp. 330–337.
[33] Wen-Tsuen Chen, Jen-Chu Liu, Hsieh-Kuan Huang, An adaptive scheme for vertical handoff in wireless overlay networks, in: Proceedings of 10th International Conference on Parallel and Distributed Systems (ICPADS), July 2004, pp. 541–548.
[34] H.J. Wang, R.H. Katz, J. Giese, Policy-enabled handoffs across heterogeneous wireless networks, in: Proceedings of 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA), 25–26 Feb. 1999, pp. 51–60.

Chun-Chuan Yang received his B.S. degree in Computer and Information Science from National Chiao-Tung University, Taiwan, in 1990 and Ph.D. degree in Computer Science from National Taiwan University in 1996. He joined the Department of Computer Science and Information Engineering, National Chi-Nan University (NCNU), Puli, Taiwan, as an Assistant Professor in 1998. Since August 2003, he has been an Associate Professor. His research area of interests includes multimedia network protocols, multimedia synchronization control, and multimedia applications.

Chun-Shian Tsai received his B.S. degree in Department of Information Management from Da-Yeh University, Taiwan, in 1998 and M.S degree in Computer Science from National Chi Nan University in 2003. He is currently working toward the Ph.D degree in Computer Science at NCNU under the guidance of Dr. Yang. His research interests include the design methodologies for multicast protocol, mobility management scheme, and the application of 802.16 network technology.

Junn-Yen Hu received his B.S. degree in Information Management from Shih Hsin University, Taiwan, in 2003 and M.S. degree in Computer Science from National Chi Nan University in 2005. Currently, he is also a Ph.D. student in Department of Computer Science of National Taiwan University. His research interests include the design of the mobility management scheme and multimedia applications.

Tzu-Chien Chuang received his B.S. and M.S degrees in Computer Science and Information Engineering from National Chi Nan University, Taiwan, in 2003 and 2005, respectively. He is now working in the Personnel Department of Tainan County Fire Bureau in Taiwan. His research interests are in the areas of wireless and wide-area computer networks, mobile computing and multimedia.