Design and Performance Study for A Mobility Management Mechanism (WMM) using Location Cache for Wireless Mesh Networks

Di-Wei Huang and **Phone Lin**[1], *Senior Member*, *IEEE*
Dept. of Computer Science & Information Engineering
National Taiwan University, Taipei 106, R.O.C.
FAX: +886-2-23628167; Email: {dwhuang@pcs., plin@}csie.ntu.edu.tw

Chai-Hien Gan, *Member*, *IEEE*
Dept. of Computer Science
National Chiao-Tung University, Hsinchu 300, R.O.C.
Email: chgan@csie.nctu.edu.tw

Abstract

Wireless Mesh Networks (WMNs) have emerged as one of the major technologies for the 4G high-speed mobile networks. In a WMN, a mesh backhaul connects the WMN with Internet, and *mesh access points* (MAPs) provide wireless network access service to *mobile stations* (MSs). The MAPs are stationary and connected through the wireless mesh links. Due to MS mobility in WMNs, Mobility Management (MM) is required to efficiently and correctly route the packets to MSs. We propose an MM mechanism named *Wireless mesh Mobility Management* (WMM). The WMM adopts the location cache approach where mesh backhaul and MAPs (referred to as *mesh nodes*; MNs) cache the MS's location information while routing the data for the MS. The MM is exercised when MNs route the packets. We implement the WMM and conduct an analytical model and simulation experiments to investigate the performance of WMM. We compare the signaling and routing cost between WMM and other existing MM protocols. Our study shows that WMM has light signaling overhead and low implementation cost.

Keywords: Location Cache; Mobility Management; Wireless Mesh Networks

---

[1]Corresponding Author

1 Introduction

Wireless Mesh Networks (WMNs) [4][6][7] have emerged as one of the major technologies for the 4G high-speed mobile networks. The WMNs provide an ubiquitous solution for wireless Internet access and MS-to-MS communication with low deployment cost. Figure 1 illustrates a general WMN architecture that comprises two kinds of fixed *mesh nodes* (MNs). The mesh backhaul is a gateway between the WMN and Internet, through which all packets are delivered between the WMN and the Internet. The *mesh access point* (MAP) provides network access service to the *mobile stations* (MSs) through the wireless access links. A wireless mesh link exists between two MNs that are located within the each other's radio coverage area. The MN location is stationary.

When an MS enters the coverage area of an MAP, the MS performs the association procedure to establish a wireless access link to the MAP [6]. This MAP is known as the *serving MAP* (SMAP) of the MS. The wireless link between the MS and the MAP can be a direct link or a relay link via other MSs. The coverage area of an MAP can be extended by relaying packets via MSs. The relay protocol exercised among MSs is out of the scope of this paper and has been studied in several previous studies (e.g., [21]). In this paper, we focus on the design of the *mobility management* (MM) protocol exercised in the fixed MNs. Before delivering the user data to an MS, the SMAP of this MS must be identified. Then the user data is sent to this SMAP through one or more MNs via the wireless mesh links. These MNs are known as the *relaying MAPs* (RMAPs). Since an MS may change the SMAP from time to time, MM is required for packet delivery to the moving MSs. Existing standards (such as IEEE 802.11 [6] and IEEE 802.16 [7]) for WMNs do not address the MM issue. The MM consists of *location management* and *handoff management*. Location management maintains the location of the current SMAP for an MS. When an MS changes its SMAP, location management updates the SMAP information for the MS. During data transmission, if the MS changes from old SMAP to new SMAP, handoff management enables the old SMAP to
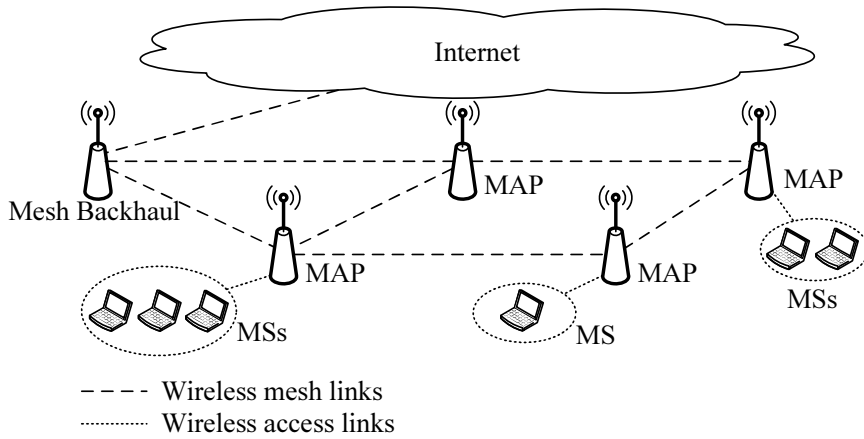
Figure 1: The WMN architecture

forward user data to the new SMAP.

Existing MM protocols for mobile networks are divided into three categories, including the ad-hoc routing protocol [21], the centralized-database MM protocol [16], and the mobile IP protocol [19]. The ad-hoc routing protocol is adopted in the *mobile ad-hoc network* (MANET) where the user data is relayed hop-by-hop by MSs, and a routing path from the source to the destination is established for routing user data. Unlike MANET, the infrastructure of WMNs is fixed (i.e., MNs are stationary). It may not be so efficient to directly apply the ad-hoc routing schemes in WMNs since most of the ad-hoc routing schemes do not consider the stationary property of WMNs. The centralized-database MM protocol (where a centralized database is maintained to store MS location information) is usually adopted in the cellular network. The service area of a cellular network is partitioned into several *location areas* (LAs). Whenever an MS moves from an LA to another, the database is accessed to update MS location information. When the size of an LA is small, high signaling cost is expected. The size of service area of an MAP may vary greatly, which depends on the radio access technology applied in WMNs. For instance, the service area of an IEEE 802.16 station can reach up to many kilometers, while an IEEE 802.11 access point can cover at most hundreds of meters. Due to the diversity of WMNs, it may not be so efficient to directly apply the centralized database for the MM protocol in WMNs. The mobile IP protocol

partitions the service area of an IP network into the home network and foreign networks. Two network entities, *home agent* (HA) in the home network and *foreign agent* (FA) in the foreign network, are responsible to tunnel user data to MSs. Similar to the centralized-database protocol, the mobile IP protocol introduces signaling overhead to inform the HA of the MS's movement. Furthermore, tunneling (between the home network and the foreign network) lengthens the routing path, which is known as the triangle routing problem [19]. Some previous studies [20] proposed optimized mobile IP protocols to overcome the triangle routing problem, where heavy extra signaling may be introduced to the network. Hence, mobile IP protocol may not be an efficient solution for WMNs.

In this paper, we propose a novel MM mechanism named *Wireless mesh Mobility Management* (WMM) for WMNs. WMM adopts the location cache approach where the MNs cache the IP address of MS's SMAP (known as MS's location information) while routing the data for the MS. The MS's location information is distributed in the MNs that have routed the packets for the MS. The MM is exercised when MNs route the packets.

We have implemented the WMM mechanism. In our implementation, we set up a WMN with four MAPs and a mesh backhaul. Each MAP is emulated by a laptop computer (running the Linux operating system) with two 802.11b WLAN cards. Our implementation is based on the Linux-2.6 OS [2] patched with the ipdivert package [1] that provides divert socket functionality for programmers to manipulate IP packets. Due to page limitation, the implementation is described in a separate report `http://pcs.csie.ntu.edu.tw/wmm/`. Our implementation is a prototype of which the major objective is to verify the functionalities of the WMM mechanism. Currently, the WMN is not widely deployed, and there is no test bed for the WMN either. The deployment of WMN will cost very much. Once the test bed is available, we will deploy the WMM mechanism in the real system, which will be considered in our future work.

The rest of the paper is organized as follows. Section 2 details the WMM mechanism.

Section 3 proposes an analytical model and simulation experiments to study the performance of WMM. Section 4 compares the signaling and routing cost of WMM and other existing MM protocols. Section 5 concludes this paper.

2   The WMM Mechanism

In the WMM mechanism, MNs are assigned fixed IP addresses. The IP addresses assigned for MSs can be done manually or by Dynamic Host Configuration Protocol (DHCP) [18]. The WMM mechanism does not require MSs to change their IP addresses for MM. An MN maintains two cache tables, the *routing table* and the *proxy table*. The routing table is used to maintain the routing paths between the MN and other MNs. The proxy table maintains the MS location information. In WMM, when an MS enters a WMN or moves from one SMAP to another MAP, the MS registers to the new SMAP. The MS location information is carried in the packet headers. When MNs route packets for an MS, the location information of the MS in proxy tables in the MNs are updated. Then the MN can correctly route the packets for MSs by referencing the proxy table and routing table. If the mesh backhaul does not cache MS location information when processing packet routing, a query procedure is executed to obtain the MS location information (to be elaborated in Section 2.3).

Several routing table maintenance protocols have been proposed in Internet or ad-hoc networks [21]. These protocols can be applied in WMM directly. In this paper, we focus on the proxy table maintenance for the MNs. As shown in Figure 2, every MN maintains an entry in the proxy table for MS, which consists of three fields: the Im field (to store MS's IP address), the Is field (to store the IP address of MS's SMAP), and the Ts field (to store the time when the MS is associated with its SMAP; also known as the "*serving timestamp*"). The serving timestamp can be obtained from the MS to ensure the non-decreasing property of the serving timestamp for the MS. We assume that all IP addresses assigned to MSs in the same WMN have the same prefix, and we can identify the WMN where the MS resides

Im Field	Is Field	Ts Field
MS's IP address	IP address of MS's SMAP	The time when the MS is associated with its SMAP

Figure 2: An entry in the proxy table

by checking the prefix of the MS's IP address. Time synchronization of MNs is required in the WMM mechanism. Existing time synchronization algorithms, such as Network Time Protocol (NTP) [17], can be used to resolve the time synchronization requirement in WMM.

We utilize the *options* field in the IP header to store the MS location information, including the IP address of MS's SMAP and MS's serving timestamp. The options field is filled or modified by MNs when they route the packets for an MS. The options field (consisting of 16 bytes) is divided into four subfields: the ISS field (to store the IP address of the sender's SMAP), the SST field (to store the sender's serving timestamp), the IRS field (to store the IP address of the receiver's SMAP), and the RST field (to store the receiver's serving timestamp). There are three WMM procedures: the registration procedure, the routing procedure, and the query procedure. Details of these procedures are described in the following subsections.

2.1 The Registration Procedure

The registration procedure is executed to register an MS to its SMAP. Suppose that MS_1 moves from its SMAP MAP_1 to another MAP MAP_2. Figure 3 illustrates the message flow for this procedure with the following steps.

Step RE1. MS_1 sends a registration request message, REREQ(MS_1's IP Address, Previous SMAP's IP Address, Selected SMAP's IP Address), to MAP_2. The previous SMAP's IP address is set to `null` if the previous SMAP is unavailable. In this example, the previous SMAP is MAP_1.

Step RE2. Upon receipt of REREQ at t_1, MAP_2 first checks whether an entry for MS_1
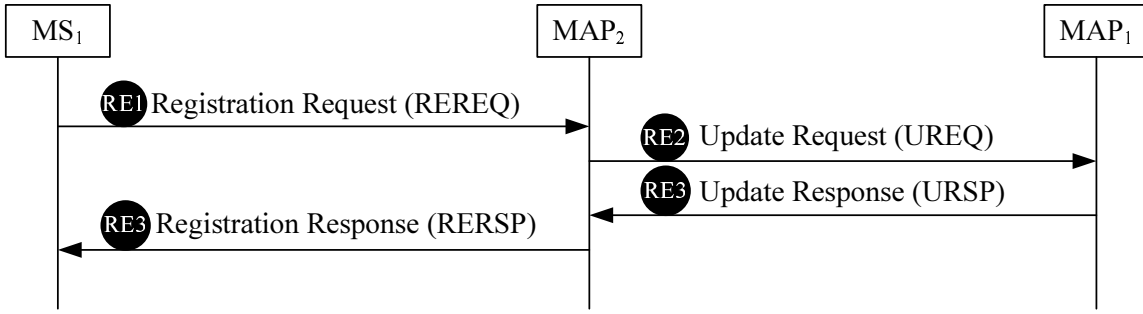
Figure 3: The message flow for the registration procedure

exists in its proxy table. If the entry exists, MAP_2 updates the entry. Otherwise, MAP_2 creates an new entry for MS_1. MS_1's entry in MAP_2's proxy table is updated as: Im is set to MS_1's IP Address; Is is set to MAP_2's IP Address; Ts is set to t_1. Then MAP_2 checks the previous SMAP's IP address carried in REREQ. If it is `null` (i.e., there is no previous SMAP for MS_1), the procedure proceeds to the next step. Otherwise, MAP_2 sends a update request message, UREQ(MS_1's IP Address, Selected SMAP's IP Address, t_1), to MAP_1. Upon receipt of UREQ, MAP_1 updates the entry for MS_1 in its proxy table as: Im is set to MS_1's IP Address; Is is set to MAP_2's IP Address; Ts is set to t_1.

Step RE3. MAP_1 responds MAP_2 an update response message, URSP. Then, MAP_2 sends a registration response message, RERSP, to MS_1, which indicates that the registration request has been completed.

Note that MS_1 may have ongoing sessions during the movement, where handoff management is required to ensure session continuity. Existing handoff management mechanisms such as enhanced IAPP [5] may be adopted in this procedure, where packets are buffered in MAP_1 and then forwarded to MAP_2. This paper concentrates the study on the location management. The details of handoff management are not included in this paper.

After Step RE3, MS_1's location information is kept in the proxy tables of both MAP_1 and MAP_2, and the location management for MS_1 is done at MAP_1 and MAP_2. For other MNs

with obsolete MS_1's location information (i.e., Is field for MS_1 stores MAP_1's IP address), the packets are first routed to MAP_1, and then MAP_1 retrieves its proxy table to forward the packets to MAP_2.

Note that there is no possibility of any loop resulted from the registration procedure. Suppose that MS_1 has the movement, MAP_1, MAP_2, ..., MAP_n, MAP_1. When MS_1 moves from MAP_n to MAP_1, the registration procedure is exercised between MS_1 and MAP_1, which updates MAP_1's proxy table. Since MAP_1 is the SMAP of MS_1, all packets can be routed to MS_1 directly through MAP_1. Hence, the loop problem does not exist in the WMM mechanism.

2.2   The Routing Procedure

The routing procedure is executed by MNs when the MNs route the packets for an MS, which consists of two parts: Location Information Synchronization and Packet Routing.

Part 1: Location Information Synchronization. In this part, the MS location information in the proxy table of the MN and that carried in the IP header of the packets are updated as the latest MS location information. Suppose that MS_1 (sender) is sending IP packets to MS_2 (receiver), where MAP_3 is one of the MNs along the routing path. Figure 4 (a) illustrates the flow chart for this part. Steps L1 and L2 update the location information for the sender (i.e., MS_1).

Step L1. Upon receipt of an IP packet, MAP_3 first checks the prefix of MS_1's IP address to determine whether MS_1 is in the WMN. If the packet is sent from Internet into the WMN (i.e., MS_1 is out of the WMN), MAP3 does not need to maintain MS_1's location information, and the procedure jumps to Step L3. Otherwise, the procedure proceeds to the next step.

Step L2. MAP_3 checks the options field in the IP header. Two cases are considered.

Case L2.I. The options field is `null`, i.e., MAP_3 is MS_1's SMAP, whose proxy table contains MS_1's current location information. MAP_3 updates the options field in the IP header: ISS is set to the Is value of MS_1's entry in MAP_3's proxy table; SST is set to the Ts value of MS_1's entry in MAP_3's proxy table; IRS is set to `null`; RST is set to the `null`.

Case L2.II. The options field is not `null`. If MS_1's entry exists in MAP_3's proxy table, MAP_3 updates MS_1's location information. Otherwise, an entry is created for MS_1 in MAP_3's proxy table. MS_1's entry in MAP_3 is set as: Im is set to MS_1's IP Address; Is is set to the ISS value in the IP header; Ts is set to the IST value in the IP header.

The following two steps (Steps L3 and L4) update the location information for the receiver (i.e., MS_2).

Step L3. MAP_3 checks the prefix of MS_2's IP address to determine whether MS_2 is in the WMN. If MS_2 is out of the WMN, the procedure exits. Otherwise, the procedure proceeds to the next step.

Step L4. This step synchronizes MS_2's location information carried in the IP header and that stored in the proxy table. Let t_t be the Ts value in MS_2's entry, and t_p be the RST value in the IP header. Without loss of generality, if MS_2's entry does not exist, $t_t = 0$, and if the RST value is `null`, $t_p = 0$. We consider three cases:

Case L4.I. $t_t < t_p$, i.e., MS_2's location information carried in the IP header is fresher than that stored in the proxy table. MS_2's location information in MAP_3's proxy table is updated as MS_2's location information carried in the IP header. MS_2's entry in MAP_3 is updated as: Im is set to MS_2's IP Address; Is is set to the RSS value in the IP header; Ts is set to the RST value in the IP header.
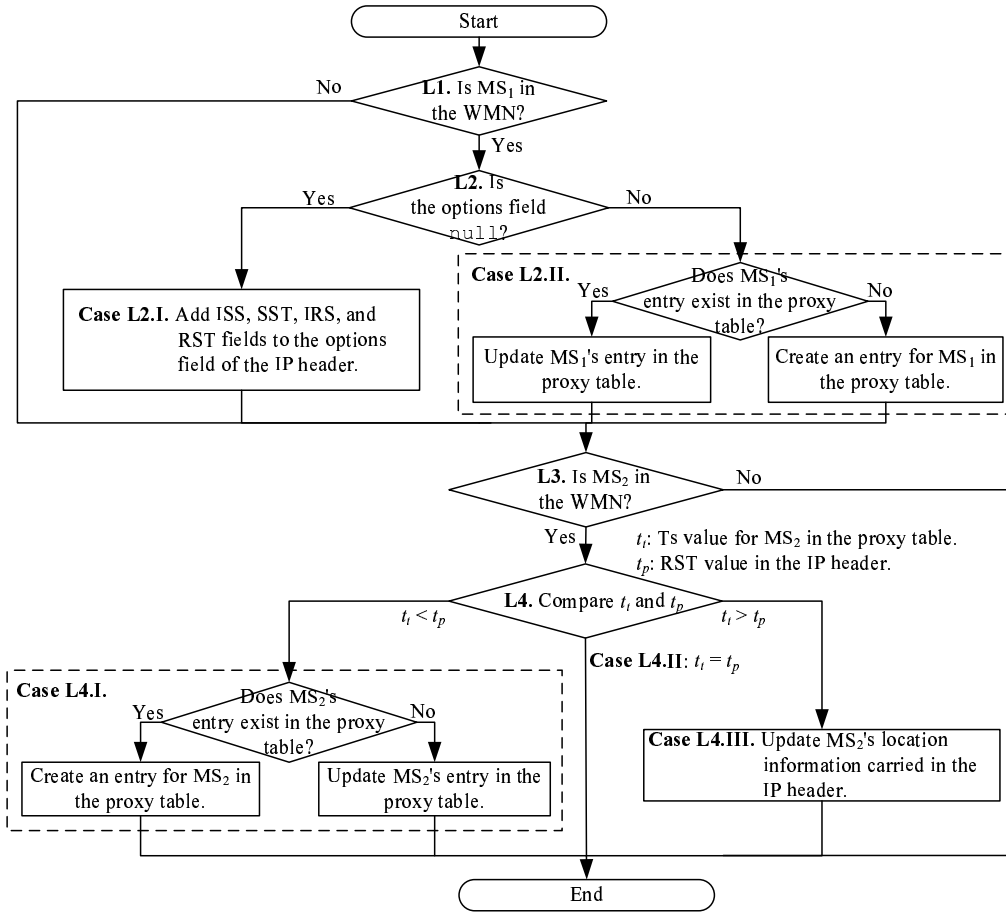
(a) Part 1: Location information synchronization
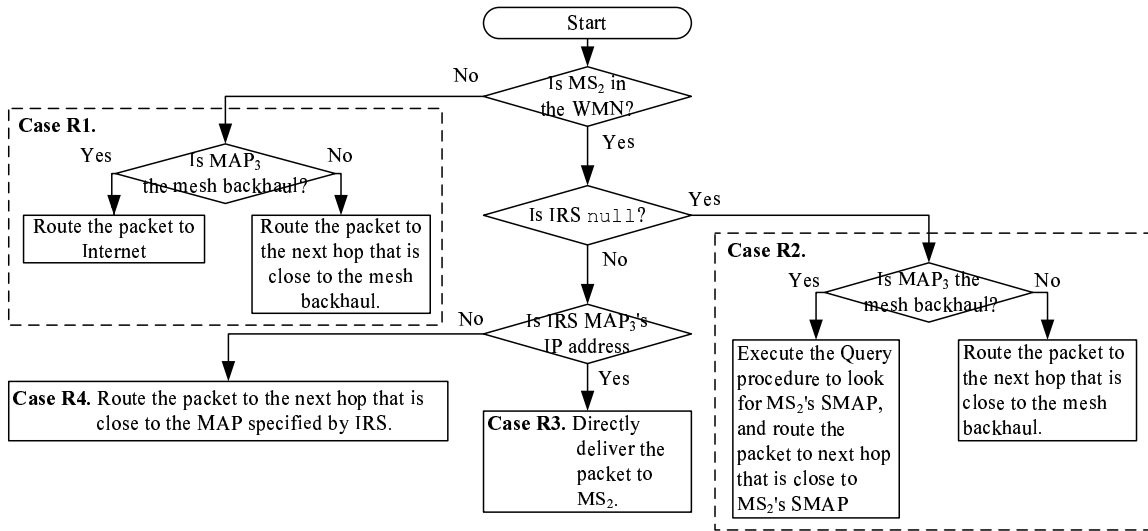
(b) Part 2: Packet routing

Figure 4: The flow chart for the routing procedure

Case L4.II. $t_t = t_p$. MS$_2$'s location information carried in the IP header is the same as that in MAP$_3$'s proxy table. The procedure does nothing.

Case L4.III. $t_t > t_p$, i.e., MS$_2$'s location information in the proxy table of MAP$_3$ is fresher than that carried in the IP header. MS$_2$'s location information carried in the IP header is filled with MS$_2$'s location information in MAP$_3$'s proxy table. The options field in the IP header is filled as: ISS is not changed; SST is not changed; IRS is set to the Is value of MS$_2$'s entry in MAP$_3$'s proxy table; RST is set to the Ts value of MS$_2$'s entry in MAP$_3$'s proxy table.

After Part 1 finishes, the sender's (i.e., MS$_1$'s) current location information is stored in MAP$_3$'s proxy table and in the IP header. MS$_2$'s location information stored in MAP$_3$'s proxy table and that carried in the IP header are synchronized.

Part 2: Packet Routing. Figure 4 (b) illustrates the flow chart for this part. In this part, MAP$_3$ routes the packets for MS$_2$ by referencing the IRS value in the IP header. MAP$_3$ determines whether MS$_2$ is in the WMN by checking the prefix of MS$_2$'s IP address, and then routes the packets by considering four cases.

Case R1. MS$_2$ is not in the WMN. If MAP$_3$ is the mesh backhaul, it simply routes the packet to Internet. Otherwise (i.e., MAP$_3$ is not the mesh backhaul), it routes the packet to the next hop (that is close to the mesh backhaul) by referencing the routing table.

Case R2. MS$_2$ is in the WMN, and the IRS field in the IP header is `null` (i.e., MS$_2$'s SMAP is unknown). If MAP$_3$ is the mesh backhaul, it exercises the query procedure (to be elaborated in the next subsection) to obtain the IP address of MS$_2$'s SMAP, and then routes the packet to the next hop that is close to MS$_2$'s SMAP. Otherwise (i.e., MAP$_3$ is not the mesh backhaul), MAP$_3$ routes the packet to the next hop that is close to the mesh backhaul by referencing the routing table.

11

Case R3. MS_2 is in the WMN, and the IRS field in the IP header specifies MAP_3's IP address (i.e., MAP_3 is MS_2's SMAP). The packet is directly delivered to MS_2.

Case R4. MS_2 is in the WMN, and the IRS field in the IP header contains the value other than MAP_3's IP address. MAP_3 references the routing table to route the packet to the next hop that is close to the MAP specified by the IRS field.

2.3 The Query Procedure

The query procedure is exercised by the mesh backhaul to obtain the IP address of receiver's SMAP when the mesh backhaul routes a packet for the receiver, and the receiver's SMAP is unknown (see Case R2 in the routing procedure). Suppose that MS_2 is the receiver of the packet. The query procedure consists of the following three steps.

Step Q1. The mesh backhaul broadcasts a route request message, RREQ(MS_2's IP Address), to all MAPs. The mesh backhaul starts a timer T_q and then expects to receive a route response message, RRES, before the timer expires.

Step Q2. Upon receipt of the RREQ message, MS_2's SMAP replies a route response message, RRES(IP Address of MS_2's SMAP, MS_2's Serving Timestamp), to the mesh backhaul.

Step Q3. If the RRES message is received before T_q expires, the mesh backhaul updates MS_2's location information carried in the IP header and that in the proxy table. After query procedure, MAP_3 can route the packet. Otherwise (i.e., T_q expires), the mesh backhaul discards the packet.

Note that the query procedure requires flooding signaling messages to all MNs in the WMN, which is a high cost operation.

3 An Analytical Model for Query Overhead

As described in Section 2.2, when the mesh backhaul routes the packet whose receiver's location information (i.e., the IP address of the receiver's SMAP) can not be determined, it exercises the query procedure to obtain the information (see Case R2). The query procedure requires flooding signaling messages to the WMN, which results in signaling overhead. This section proposes an analytical model and simulation experiments to study this performance issue.

We classify the traffic in a WMN into two categories: Internet and intranet sessions. The Internet session involves an MS and a server (or a host) out of the WMN, which are initiated by the MS. The packets for Internet sessions must be routed through the mesh backhaul, and the MS's location information in the mesh backhaul's proxy table is updated. The intranet session involves two MSs in the same WMN.

Consider the timing diagram in Figure 5. Suppose that MS_0 enters the WMN at t_0. Let x be the time period between t_0 and the time when MS_0 originates the 1st Internet session. Suppose that the Internet session arrivals originated by MS_0 form a Poisson process with rate λ. Then, with the memoryless property of the exponential distribution, we have the density function $f_x(x)$ for x as

$$f_x(x) = \lambda e^{-\lambda x}.$$

Suppose that when MS_0 enters the WMN, there are another N MSs. We assume that the N MSs are identical, and each MS initiates intranet sessions toward MS_0 with probability γ. Let N' ($0 \leq N' \leq N$) be the number of MSs (that initiate intranet sessions toward MS_0). Without loss of generality, we assume that the N' MSs are MS_1, MS_2, ..., $MS_{N'}$. Let y_k be the time period between t_0 and the time when MS_k (where $1 \leq k \leq N'$) originates the 1st intranet session toward MS_0, where y_k is assumed to have a general distribution with the density function $f_y(\cdot)$ and the distribution function $F_y(\cdot)$.
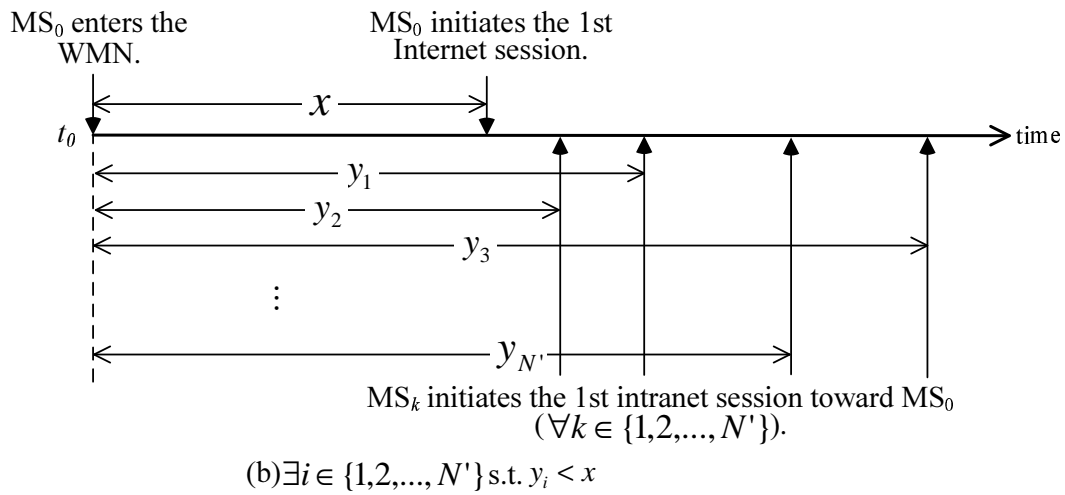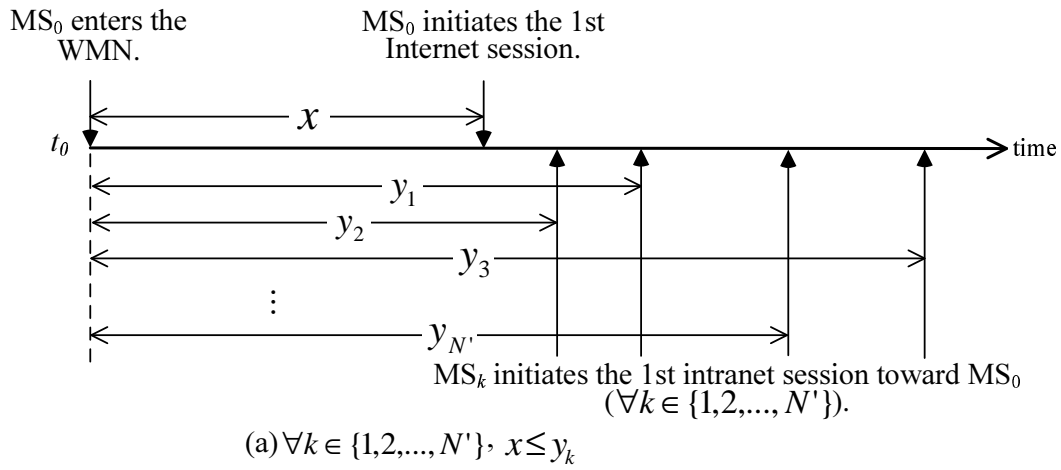
13

(a) $\forall k \in \{1,2,...,N'\}$, $x \le y_k$

(b) $\exists i \in \{1,2,...,N'\}$ s.t. $y_i < x$

Figure 5: The timing diagram

14

Let P_q be the probability that the query procedure is invoked by the mesh backhaul to obtain MS_0's location information. As shown in Figure 5, two cases are considered to derive the P_q probability:

Case 1. For all $k \in \{1, 2, ..., N'\}$, $x \leq y_k$ (see Figure 5 (a)). In this case, MS_0 initiates the 1st Internet session at $t_0 + x$, where the mesh backhaul creates an entry to store MS_0's location information. After $t_0 + x$, if there are packets (either for Internet sessions or intranet sessions) to be routed to MS_0, these packets can be correctly routed to MS_0 without invoking the query procedure.

Case 2. There exists $i \in \{1, 2, ..., N'\}$, such that $y_i \leq x$ (see Figure 5 (b)). In this case, at least one MS_i ($1 \leq i \leq N'$) initiates the 1st intranet session toward MS_0 at $t_0 + y_i$ before MS_0 initiates the 1st Internet session at $t_0 + x$. At $t_0 + y_i$, if any of the MNs along the routing path between MS_i's SMAP and the mesh backhaul stores MS_0's location information, then the query procedure is not invoked during the packet transmission for the intranet session from MS_i to MS_0.

Let A be the probability that Case 1 occurs, and B be the probability that given Case 2, the routing MNs contain MS_0's location information. Then we have the P_q probability as

$$P_q = 1 - A - B.$$

Probability B is highly dependent on the network topology and the relative positions of the MSs. The analysis for B is too complicated. In this study, we derive an upper bound $\tilde{P}_q$ for P_q, that is

$$\tilde{P}_q = 1 - A \geq P_q. \tag{1}$$

The A probability is derived as follows:

$$
\begin{aligned}
A &= \sum_{j=0}^{N} \Pr[\forall k \in \{1, 2, ..., N'\}, x \le y_k | N' = j] \\
&= \sum_{j=0}^{N} \Pr[\forall k \in \{1, 2, ..., j\}, x \le y_k] \Pr[N' = j] \\
&= \sum_{j=0}^{N} \left\{ \int_{x=0}^{\infty} \left[\prod_{k=1}^{j} \int_{y_k=x}^{\infty} f_y(y_k) dy_k \right] f_x(x) dx \right\} \binom{N}{j} \gamma^j (1-\gamma)^{N-j} \\
&= \sum_{j=0}^{N} \left\{ \int_{x=0}^{\infty} \left[\prod_{k=1}^{j} \int_{y_k=x}^{\infty} f_y(y_k) dy_k \right] \lambda e^{-\lambda x} dx \right\} \binom{N}{j} \gamma^j (1-\gamma)^{N-j} \\
&= \sum_{j=0}^{N} \left\{ \lambda \int_{x=0}^{\infty} [1 - F_y(x)]^j e^{-\lambda x} dx \right\} \binom{N}{j} \gamma^j (1-\gamma)^{N-j}. \quad (2)
\end{aligned}
$$

Let $\overline{F}_y^j(y) = [1 - F_y(y)]^j$ and $f_y^*(s)$ be the Laplace transform of $\overline{F}_y^j(y)$. Then (2) is rewritten as

$$
A = \lambda \sum_{j=0}^{N} f_y^*(\lambda) \binom{N}{j} \gamma^j (1-\gamma)^{N-j}. \quad (3)
$$

Apply (3) into (1), and $\tilde{P}_q$ is expressed as

$$
\tilde{P}_q = 1 - \lambda \sum_{j=0}^{N} f_y^*(\lambda) \binom{N}{j} \gamma^j (1-\gamma)^{N-j}. \quad (4)
$$

Our analysis can apply any y_k distribution whose $f_y^*(s)$ exists. Here, we take the exponential y_k distribution (with mean $1/\eta$) as an example. Then, we have

$$
\overline{F}_y^j(y) = \left[1 - \left(1 - e^{-\eta y} \right) \right]^j = e^{-j\eta y} \quad (5)
$$

and

$$
f_y^*(s) = \int_0^{\infty} e^{-j\eta y} e^{-sy} dy = \frac{1}{j\eta + s} \quad (6)
$$

Applying (6) into (4), we have

$$
\tilde{P}_q = 1 - \lambda \sum_{j=0}^{N} \frac{1}{j\eta + \lambda} \binom{N}{j} \gamma^j (1-\gamma)^{N-j}. \quad (7)
$$

This study also conducts simulation experiments to investigate the P_q performance. We adopt the discrete event-driven approach in our simulation, which has been widely used to simulate the mobile communication networks in several studies [9][10][11][13][14]. Following the standard [7], a WMN is modeled as a regular hexagonal topology. Each hexagon represents the coverage area of an MAP. In our simulation, the WMN consists of 61 MNs (i.e., 1 mesh backhaul + 60 MAPs) and 1000 MSs (i.e., $N = 1000$). The mesh backhaul is located at the center of the WMN. The movement of an MS follows a 2-D random walk model [3], where an MS resides in an MAP's coverage area for a period and then moves to one of its neighboring MAPs with the same probability 1/6. We denote the MS resident time in an MAP's coverage area as z. Due to the page limitation, the details of the simulation are omitted in this paper. Figure 6 compares the results (see the solid curves) of the analytical model against that (see the dashed curves) of the the simulation, where the parameter setups will be elaborated later. The figures indicate that P_q of the simulation model is properly bounded by $\tilde{P}_q$. The trends of P_q and $\tilde{P}_q$ curves are the same.

In the following, we run simulation experiments to investigate the P_q performance for WMM, where we assume x (i.e., the time period between t_0 and the time when MS_0 originates the 1st Internet session) and y_k (i.e., the time period between t_0 and the time when MS_k, $1 \le k \le N'$, originates the 1st intranet session toward MS_0) have exponential distributions with means $1/\lambda$ and $1/\eta$, respectively. In our experiments, we assume that $\eta \ll \lambda$ (or $1/\eta \gg 1/\lambda$). The reason is that, typically, an MS is likely to initiate an Internet session in the early time. For example, an MS may register to a SIP Proxy server immediately after it enters WMN. On the other hand, it may take time for MS_k's application to detect that MS_0 is in the WMN before it initiates an intranet session toward MS_0. The MS resident time z is assumed to be Gamma distributed with mean $1/\omega$ and variance υ_z, respectively. The Gamma distribution has been widely adopted to simulate MS moving behavior in the real mobile networks in several studies [8][9][12][15]. The input parameters η and ω are normalized by λ. For example, if $1/\lambda = 1$ second, then $\omega = 0.1\lambda$ means that the expected
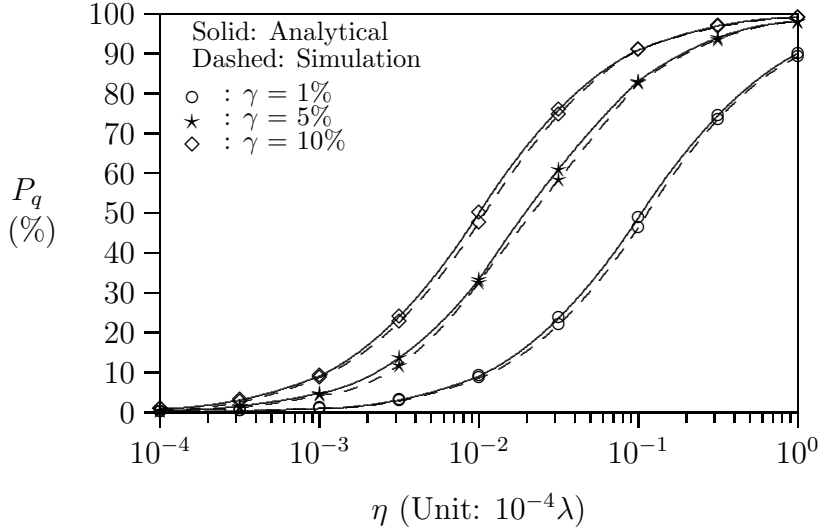
Figure 6: Validation of the simulation and the analytic models ($v_z = \frac{1}{\omega^2}$; $\omega = 0.1\lambda$; $N = 1000$)

MAP residence time is 10 seconds. The impacts of the input parameters are discussed below.

Effects of η. Figure 6 plots P_q as an increasing function of η, where we set $\omega = 0.1\lambda$, $N = 1000$, and $v_z = 1/\omega^2$ (i.e., exponential MAP residence time). As η increases, other MSs initiate their 1st intranet session to MS_0 earlier (i.e., $1/\eta$ is smaller). It is more likely that an MS initiates an intranet session to MS_0 before MS_0 initiates the 1st Internet session, and the mesh backhaul may not contain MS_0's location information when it routes the intranet session for MS_0. Therefore, we observe the larger P_q values when η increases.

Effects of γ. Figure 6 also studies the effects of γ, where γ is set to 1%, 5%, and 10%. The figure shows that with lager γ, the P_q performance for WMM drops. When γ increases, there are more MSs that initiate an intranet session toward MS_0. It is more likely that an MS initiates an intranet session to MS_0 before MS_0 initiates the 1st Internet session. The mesh backhaul has worse chance to contain MS_0's location information when it routes the intranet session for MS_0.

Effects of ω. Figure 7 studies the effects of ω, where $\eta = 0.001\lambda$, $v_z = 1/\omega^2$ (i.e., exponen-
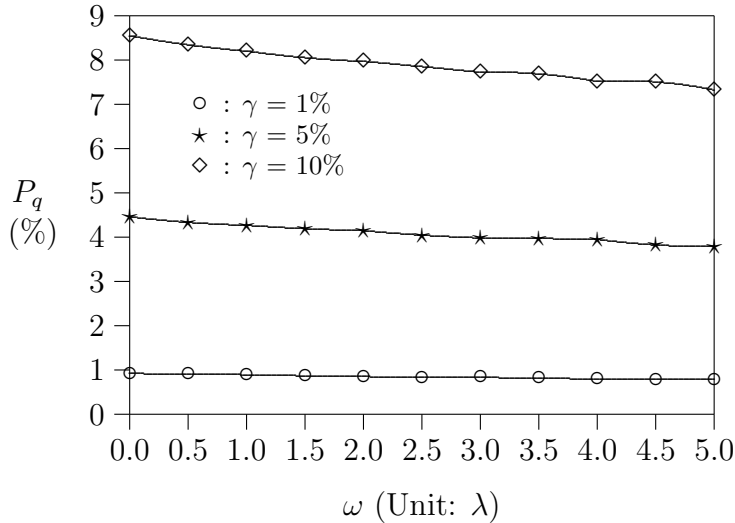
18

Figure 7: Effects of ω ($\eta = 10^{-3}\lambda$; $v_z = \frac{1}{\omega^2}$; $N = 1000$)

tial MAP residence time), and $N = 1000$. As ω increases, the P_q values slightly drop. With the higher MS mobility, through the registration procedure, the MS's location information are more likely to be cached in the MAPs, which reduces the possibility to invoke the query procedure. Therefore, we observe that P_q decreases as ω increases.

Effects of v_z. Figure 8 studies the effects of v_z, where $\eta = 0.001\lambda$, $\omega = 0.1\lambda$, and $N = 1000$. The figure shows that with lager v_z, the P_q values drop. This is due to that as v_z becomes large, more small z values are observed. Therefore, the higher MS mobility is expected. Similar to the effects of ω (see Figure 7), the MS's location information are more likely to be cached in the MAPs. We observe that WMM functions better (i.e., smaller P_q is observed) as v_z increases.

4 Comparison between WMM and Existing MM Protocols

This section compares the WMM mechanism with other existing MM protocols (including ad-hoc routing protocol, centralized-database MM protocol, and mobile IP protocol) in terms of location update, location tracking, and packet routing cost. More nodes (involved in
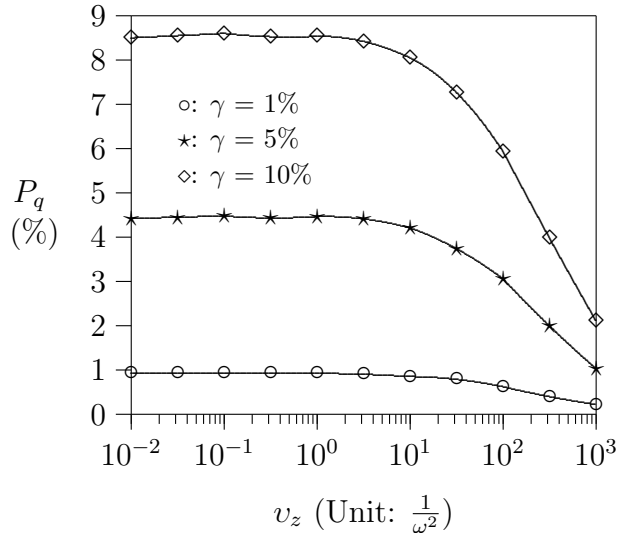
19

Figure 8: Effects of variance v_z ($\eta = 10^{-3}\lambda$; $\omega = 0.1\lambda$; $N = 1000$)

a location update operation and a location tracking operation) result in more signaling messages replicated within the WMN (i.e., more signaling traffic) and longer routing path for signaling message delivery (i.e., the possibility for successful message delivery decreases). Furthermore, the total computation overhead for each MN and MS to process location update and location tracking messages in the WMN increases. Therefore, we define the location update cost C_u and the location tracking cost C_t as the average number of MNs and MSs that exchange signaling messages for location update operation (executed when an MS changes its SMAP) and location tracking operation (executed when a session is initiated toward an MS), respectively. When there are more nodes involved in routing a packet (i.e., longer routing path), the packet delivery latency increases, and the possibility for successful packet delivery decreases. Therefore, we define the packet routing cost C_r as the average number of MNs or MSs that route a packet to a destination MS. We consider a WMN consisting of M MNs and N MSs. The C_u, C_t, and C_r costs for WMM and other three existing MM protocols are compared in the following subsections. The notations used in this comparison is listed below.

- M (N): the number of MNs (MSs) in the WMN

20

- $\bar{R}$: the average number of MNs in the routing path between two MSs or between an MS and a centralized node (e.g., a centralized database or an HA)

- P_q: the probability that the query procedure in WMM mechanism is invoked for an MS

- r: the average number of sessions initiated toward an MS in the WMN

Usually, not all MNs are involved in routing packets between two arbitrary nodes. Hence, we have $M \geq \bar{R}$. In this study, we consider WMNs with large scale (i.e., M and $\bar{R}$ are large numbers).

4.1 Signaling and Routing Cost for WMM

In WMM, when an MS changes its SMAP, location update is done through the registration procedure and the routing procedure. The registration procedure is executed among an MS, the MS's current SMAP, and the MS's pervious SMAP. We consider three situations:

S1. The MS enters the WMN and then is powered on. The MS only communicates with its SMAP. There are two nodes involved in signaling message exchange for the registration procedure. The cost for the registration is 2.

S2. The MS is powered on, and moves from the old SMAP to the new SMAP. There are three nodes involved in the signaling message exchange for the registration procedure. The cost for the registration is 3.

S3. The MS is switched off at $SMAP_1$, and then powered on at $SMAP_2$. This is taken as a new registration to $SMAP_2$. There are two nodes involved in the signaling message exchange of the registration procedure. The cost for the registration is 2.

Based on the above discussion, the cost for the signaling message exchange of the registration procedure is less than or equal to 3. The routing procedure is done by an MN while it routes

a packet for the MS. No signaling messages are required for the routing procedure. Thus, we have $C_u \leq 3$ for WMM.

When a session is initiated toward an MS, location tracking is processed through the routing procedure and the query procedure. As described above, the routing procedure does not incur signaling cost. On the other hand, the query procedure requires flooding signaling messages to all MNs. The number of nodes involved in the query procedure equals to the total number of MNs (i.e., M). However, the query procedure may not be invoked for the MS during the time when the MS stays in the WMN. Actually, the query procedure is executed with probability P_q for an MS (see Section 3), and it is invoked at most once for the MS. Let r ($r > 0$) be the number of sessions initiated toward an MS during the time when the MS stays in the WMN. Consequently, C_t for WMM can be estimated as $\dfrac{M \cdot P_q}{r}$.

Suppose that MS_1 is sending packets to MS_2, where MAP_1 and MAP_2 are SMAPs of MS_1 and MS_2, respectively. Let MAP'_2 be MS_2's previous SMAP. Three cases are considered to count C_r for WMM:

Case 1. MAP_1's proxy table contains MS_2's current location information. The packets can be routed directly to MS_2, and C_r is $\bar{R}$.

Case 2. MAP_1's proxy table contains obsolete MS_2's location information. The packets are first routed to MAP'_2, and then MAP'_2 routes the packets to MAP_2 through other MAPs, where routing path between MAP'_2 and MAP_2 may be a direct link or polygon links. Let $\bar{r}_1$ be the average number of MNs along the routing path between MAP_1 and MAP'_2 and $\bar{r}_2$ be the average number of MNs along the routing path between MAP'_2 and MAP_2. In this case, we have $C_r = \bar{r}_1 + \bar{r}_2 > \bar{R}$.

Case 3. MS_2's entry does not exist in MAP_1's proxy table. The packets will be routed to the mesh backhaul. If an MN along the routing path between MAP_1 and the mesh backhaul contains MS_2's location information, then the packet can be routed to MS_2.

Suppose that the average number of MNs in the routing path between MAP_1 and mesh backhaul is r_3, and the average number of MNs in the routing path between mesh backhaul and MAP_2 is r_4. In this case, $C_r = r_3 + r_4 > \bar{R}$.

Let the probability that Case 1 occurs be β_1, the probability that Case 2 occurs be β_2, the probability that Case 3 occurs be β_3. We have

$$\beta_1 + \beta_2 + \beta_3 = 1. \tag{8}$$

In most of the cases in the real system, communications between two MSs are bidirectional (i.e., both MS_1 and MS_2 exchange packets with each other). Once MS_1 and MS_2 exchange packets with each other, MS_1's SMAP caches MS_2's current location information and vice versa, and the prolonged routing path (incurring in Case 2 or 3) will be changed to a direct routing path. At this moment, the C_r cost for Case 2 or Case 3 is $\bar{R}$. Suppose that MS_1 sends the first packet at time t_0, MS_1 and MS_2 start bidirectional communication at time t_1, and the communication ends at time t_2, where $t_2 > t_1 > t_0$. Therefore, the C_r cost can be estimated as

$$C_r = \beta_1 \bar{R} + \beta_2 \left[\left(\frac{t_1 - t_0}{t_2 - t_0} \right) (\bar{r}_1 + \bar{r}_2) + \left(\frac{t_2 - t_1}{t_2 - t_0} \right) \bar{R} \right] + \beta_3 \left[\left(\frac{t_1 - t_0}{t_2 - t_0} \right) (\bar{r}_3 + \bar{r}_4) + \left(\frac{t_2 - t_1}{t_2 - t_0} \right) \bar{R} \right]. \tag{9}$$

Since in most of applications (e.g., TCP session), the bidirectional communication usually starts after MS_2 receives the first packet from MS_1, we have $t_1 \approx t_0$, i.e, $\frac{t_1 - t_0}{t_2 - t_0} \approx 0$ and $\frac{t_2 - t_1}{t_2 - t_0} \approx 1$. With (8), C_r in (9) approximates to $\bar{R}$.

4.2 Signaling and Routing Cost of Ad-hoc Routing Protocol

Two basic approaches, proactive (also known as table-driven) and reactive (also known as demand-driven) are proposed for the ad-hoc routing protocol [21]. In the proactive approach, an MS maintains a routing table to store all routing paths between the MS and other MSs. Location update is done by notifying all MNs and MSs of the MS's movement, and we have

$C_u = M + N$ for the proactive ad-hoc routing protocol. In the proactive approach, when an MS routes a packet to the destination MS, it references its own routing table, and no signaling messages are required for location tracking. Thus, $C_t = 0$ for the proactive ad-hoc routing protocol. Furthermore, in the proactive ad-hoc routing protocol, since the routing tables always contain current location information for MSs, packets can be routed directly to the destination. We have $C_r = \bar{R}$ for the proactive ad-hoc routing protocol.

In the reactive approach, the MS discovers routes when it has packets to be sent. No location update operation is executed in this approach, and we have $C_u = 0$ for the reactive ad-hoc routing protocol. When an MS has packets to be sent to the destination MS, it obtains a route through flooding signaling messages to the WMN, and the location tracking is done at the same time. Some reactive approaches for ad-hoc routing protocols (e.g., Ad hoc On-demand Distance Vector; AODV [21]) can cache the old route until the old route breaks. Let n_s be the average number of sessions between any two MSs, MS_0 and MS_1, within a time period where MS_0 and MS_1 do not change their SMAP. To initiate the first session, MS_0 or MS_1 floods signaling messages to the WMN and obtains a routing path between MS_0 and MS_1. Since the total number of nodes in the WMN is $M + N$, the C_t cost for the first session is $M + N$. For the others of the n_s sessions, Since MS_0 (or MS_1) has cached the routing information after the first session, there is no need to flood any signaling messages for the establishment of other sessions. The average signaling cost for one session can be estimated as $C_t = \dfrac{M + N}{n_s}$. Typically, when the mobility of MSs is high, the number n_s is a small number.

To summarize, the proactive approach introduces extra signaling overhead for routing table maintenance especially when MS mobility is high. The reactive approach not only introduces extra signaling overhead but also spends time to establish the route before user data is delivered, which significantly delays user data transmission.

4.3 Signaling and Routing Cost of Centralized-database MM Protocol

In centralized-database MM protocol, a centralized MM database is maintained to store the location information for all MSs. Whenever an MS moves from an LA to another, a registration procedure is triggered for location update. The registration procedure is executed between the MS and the database to update the LA ID stored in the database. Thus, we have $C_u = \bar{R}$ for the centralized-database MM protocol.

When an MS has packets to be sent to the destination MS, it queries the database for the destination MS's location information, where the signaling messages for location tracking are exchanged between the MS and the database. Thus, we have $C_t = \bar{R}$ for the centralized-database MM protocol.

In the centralized-database MM protocol, the destination MS's current location information is stored in the centralized database. The packets can be routed correctly to the destination MS. We have $C_r = \bar{R}$ for the centralized-database MM protocol.

4.4 Signaling and Routing Cost of Mobile IP Protocol

In mobile IP protocol, the HA in the home network and the FA in the foreign network are responsible to tunnel packets for MSs. When an MS moves from the home network to the foreign network, a registration procedure is triggered for location update, which is executed between the MS, the FA, and the HA to inform the FA and HA of the MS's movement. Typically, the MS is close to the FA, and we can omit the signaling and routing cost between the MS and the FA. Thus, we have $C_u = \bar{R}$ for the mobile IP protocol.

When MNs route packets to an MS, the packets are first routed to the HA, and then the HA routes the packets to the destination MS directly (if the destination MS is in its home network) or delivers the packets to the destination MS by tunneling them from the HA to the FA (if the destination MS is in the foreign network). There is no signaling messages

required for location tracking, and we have $C_t = 0$ for the mobile IP protocol.

In mobile IP protocol, the packets are always routed to the HA and then to the destination (i.e., the triangle routing problem), the C_r cost for the mobile IP protocol can be estimated as $2\bar{R}$.

The previous study [20] proposed route optimization for mobile IP protocol to overcome triangle routing. Suppose that a corresponding node MS_1 is sending packets toward MS_2 in the foreign network. Initially, with the base mobile IP protocol, the packets are routed to the HA. Then the HA sends a Binding Update message, which contains the IP address of the MS_2's current FA, to MS_1, and all packets of MS_1 will be routed to MS_2's current FA. Routing for the Binding Update message introduces signaling overhead for location tracking, which is estimated as $C_t = \bar{R}$ for mobile IP with route optimization. The registration procedure for mobile IP with route optimization is the same as that for base mobile IP, and the C_u costs are the same, which is $C_u = \bar{R}$. With route optimization, the C_r cost is estimated as $C_r = \bar{R}$.

4.5 Comparison

Table 1 lists the C_u cost and the C_t cost for the four existing MM protocols and the proposed WMM mechanism. As shown in Table 1, the C_u cost for WMM is constant and is lower than that for the proactive ad-hoc routing protocol, the centralized-database MM protocol, and the mobile IP protocol. The C_u cost for WMM is slightly higher than that for the reactive ad-hoc routing protocol. However, the C_t for WMM is much lower than that for the reactive ad-hoc routing protocol.

The C_t cost for WMM is estimated as $\dfrac{M \cdot P_q}{r}$, where $P_q \leq 1$. Typically, the service area of a WMN is not huge, and the number (i.e., M) is not large number. Furthermore, when the traffic load to a MS increases (i.e., more sessions are initiated to an MS; i.e., r increases), we gain better C_t performance for WMM. From this analysis, we have that the

Table 1: Comparison between WMM and other MM mechanisms

	The C_u Cost	The C_t Cost	The C_r Cost
Proactive Ad-hoc Routing Protocol	$M + N$	0	$\bar{R}$
Reactive Ad-hoc Routing Protocol	0	$\dfrac{M + N}{n_s}$	$\bar{R}$
Centralized-database MM Protocol	$\bar{R}$	$\bar{R}$	$\bar{R}$
Base Mobile IP Protocol	$\bar{R}$	0	$2\bar{R}$
Mobile IP Protocol with Route Optimization	$\bar{R}$	$\bar{R}$	$\bar{R}$
WMM Mechanism	≤ 3	$\dfrac{M \cdot P_q}{r}$	$\approx \bar{R}$

WMM mechanism does not incur heavy C_t cost. Obviously, compared with the reactive ad-hoc routing protocol (whose C_t is estimated as $\dfrac{M + N}{n_s}$), the centralized-database MM protocol (whose C_t is estimated as $\bar{R}$), and the mobile IP with route optimization (whose C_t is estimated as $\bar{R}$, WMM has the lower C_t cost. The C_t for WMM is slightly higher than that for the proactive ad-hoc routing protocol and the base mobile IP protocol. However, the WMM significantly outperforms the proactive ad-hoc routing protocol and the base mobile IP protocol in terms of the C_u cost.

The C_r cost for WMM is obviously lower than that for the base mobile IP protocol, whose C_r cost is estimated as $2\bar{R}$. The C_r cost for WMM approaches that for the proactive ad-hoc routing protocol, the reactive ad-hoc protocol, the centralized-database MM protocol, and the mobile IP with route optimization. However, as mentioned above, the ad-hoc routing protocols, the centralized-database protocol, and the mobile IP with route optimization incurs much heavier signaling overhead (i.e., C_u and C_t cost) than WMM does. To summarize, the proposed WMM mechanism is capable of correctly and efficiently routing packets for MSs with lighter overhead than the existing MM protocols do.

5   Concluding Remark

This paper designed a novel MM protocol, the WMM mechanism, for WMNs by capturing the characteristics of WMNs. In the WMM mechanism, location caches to cache MSs'

location information are added in the MAPs so that the network can more efficiently (i.e., fast and low signaling cost) routes packets to mobile users. The fields in IP header are utilized to carry MSs' location information. Location update can be done at the same time when the MN routes packets for MSs, and the signaling cost for location update is reduced. The MSs' location information is distributed within the WMN. We have implemented a prototype of the WMM mechanism in the real system.

If the MS's location information can not be determined, the query procedure is executed to find the MS's location, where signaling cost may incur to the network. We conducted an analytical model and simulation experiments to study this performance issue. Our study shows that

- When an MS enters the WMN, if the 1st intranet session is initiated toward the MS earlier, the P_q is higher. On the other hand, if the MS initiates its 1st Internet session earlier, the P_q is lower.

- With higher MS mobility or higher variance of the MAP residence time, the WMM gains the better P_q performance (i.e, P_q is lower).

In the end of the paper, we made a comparison between WMM and other existing MM protocols (including ad-hoc routing protocol, centralized-database MM protocol, and mobile IP protocol). Our study concluded that the WMM mechanism can provide correct and efficient packet routing for MSs with lighter signaling and routing cost than the existing MM protocols do.

Acknowledgement

reviewers for their valuable comments. Their efforts have significantly improved the quality of our paper.

References

[1] Divert Sockets for Linux. [Online]. Available: http://sourceforge.net/projects/ipdivert/.

[2] The Linux Kernel Archives. [Online]. Available: http://www.kernel.org/.

[3] Akyildiz, I. F., Lin, Y.-B., Lai, W.-R., and Chen, R.-J. A New Random Walk Model for PCS Networks. *IEEE Journal on Selected Areas in Communications*, 18(7):1254–1260, July 2000.

[4] Akyildiz, I. F., Wang, X., and Wang W. Wireless Mesh Networks: A Survey. *Computer Networks Journal, Elsevier*, 47(4):445–487, March 2005.

[5] Chou, C.-T. and Shin, K. G. An Enhanced Inter-Access Point Protocol for Uniform Intra and Intersubnet Handoffs. *IEEE Journal on Mobile Computing*, 14(4):321–334, July-August 2005.

[6] IEEE. IEEE Std 802.11-1997 Information Technology- telecommunications And Information exchange Between Systems-Local And Metropolitan Area Networks-specific Requirements-part 11: Wireless LAN Medium Access Control (MAC) And Physical Layer (PHY) Specifications. *IEEE Std 802.11-1997*, November 1997.

[7] IEEE. IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. *IEEE Std 802.16-2004 (Revision of IEEE Std 802.16-2001)*, May 2004.

[8] Lin, P. and Lin, Y.-B. Channel Allocation for GPRS. *IEEE Transactions on Vehicular Technology*, 50(2):375–387, March 2001.

[9] Lin, P. and Tu, G.-H. An Improved GGSN Failure Restoration Mechanism for UMTS. *ACM Wireless Networks*, 12(1):91–103, February 2006.

[10] Lin, P., Chang, H.-M., Fang, Y., and Cheng, S.-M. HISNs: Distributed Gateways for Application-level Integration of Heterogenous Wireless Networks. *ACM Wireless Networks*, Accepted for publication.

[11] Lin, P., Lai, W.-R., and Gan, C.-H. Modeling Opportunity Driven Multiple Access in UMTS. *IEEE Transactions on Wireless Communications*, 3(5):1669–1677, September 2004.

[12] Lin, P., Lin, Y.-B., and Chlamtac, I. Modeling Frame Synchronization for UMTS High-Speed Downlink Packet Access. *IEEE Transactions on Vehicular Technology*, 50(1):132–141, January 2003.

[13] Lin, P., Lin, Y.-B., and Chlamtac, I. Module Count-Based Overflow Control Scheme for UMTS High-Speed Downlink Packet Access. *IEEE Transactions on Vehicular Technology*, 53(2):425–432, March 2004.

[14] Lin, P., Lin, Y.-B., and Chlamtac, I. Overflow Control for UMTS High-Speed Downlink Packet Access. *IEEE Transactions on Wireless Communications*, 3(2):524–532, March 2004.

[15] Lin, P., Lin, Y.-B., Gan, C.-H., and Jeng, J.-Y. Credit allocation for UMTS prepaid service. *IEEE Transactions on Vehicular Technology*, 54(6):306–316, November 2005.

[16] Lin, Y.-B. and Chlamtac, I. *Wireless and Mobile Network Architectures*. John Wiley & Sons, 2001.

[17] Mills, D.L. Internet time synchronization: the network time protocol. *IEEE Transactions on Communications*, 39(10):1482–1493, October 1991.

[18] Network Working Group. Dynamic Host Configuration Protocol. *Internet RFC 2131/STD*, March 1997.

[19] Perkins, C. E. Mobile Networking through Mobile IP. *IEEE Internet Computing*, 2(1):58–69, January 1998.

[20] Perkins, C. E. and Johnson, D. B. Route Optimization for Mobile IP. *Cluster Computing*, 1(2):161–176, June 1998.

[21] Royer, E. M. and Toh, C.-K. A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks. *IEEE Personal Communications*, 6(2):46–55, April 1999.