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

Di-Wei Huang, Phone Lin, Senior Member, IEEE, and Chai-Hien Gan, Member, IEEE

Abstract—Wireless Mesh Networks (WMNs) have emerged as one of the major technologies for 4G high-speed mobile networks. In a WMN, a mesh backhaul connects the WMN with the 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.

Index Terms—Location cache, mobility management, wireless mesh networks.

1 INTRODUCTION

Wireless Mesh Networks (WMNs) have emerged as one of the major technologies for 4G high-speed mobile networks. The WMNs provide a ubiquitous solution for wireless Internet access and MS-to-MS communication with low deployment cost. Fig. 1 illustrates a general WMN architecture that comprises two kinds of fixed mesh nodes (MNs). The mesh backbone 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 each other’s radio coverage area. The MN location is stationary.

When an MS enters the coverage area of a MAP, the MS performs the association procedure to establish a wireless access link to the MAP [2]. 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 a 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 (for example, [4]). 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 [2] and IEEE 802.16 [3]) 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 forward user data to the new SMAP.

Existing MM protocols for mobile networks are divided into three categories, including the ad hoc routing protocol [4], the centralized-database MM protocol [5], and the mobile Internet Protocol (IP) [6]. 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 (that is, 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 one LA to another, the database is accessed to update the MS location information. When the size of an LA is small, high signaling cost is expected. The size of the service area of a MAP may vary greatly, which depends on the radio access technology applied in WMNs. For instance, the service area of an IEEE
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 overhead 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) [10]. 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 the Internet or ad hoc networks [4]. These protocols can be applied in WMM directly. In this paper, we focus on the proxy table maintenance for the MNs. As shown in Fig. 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 time stamp”). The serving time stamp can be obtained from the MS to ensure the nondecreasing property of the serving time stamp 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 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) [11] 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 time stamp. 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 time stamp), the IRS field (to store the IP address of the receiver’s SMAP), and the RST field (to store the receiver’s serving time stamp). There are three WMM procedures: the registration procedure, the routing procedure, and the query procedure. Details of these procedures are described in the following sections.
Step RE1. MS1 sends a registration request message, REREQ (MS1’s IP Address, Previous SMAP’s IP Address, Selected SMAP’s IP Address) to MAP2. The previous SMAP’s IP address is set to null if the previous SMAP is unavailable. In this example, the previous SMAP is MAP1.

Step RE2. Upon receipt of REREQ at t1, MAP2 first checks whether an entry for MS1 exists in its proxy table. If the entry exists, MAP2 updates the entry. Otherwise, MAP2 creates a new entry for MS1. MS1’s entry in MAP2’s proxy table is updated as follows: Im is set to MS1’s IP Address, Is is set to MAP2’s IP Address, and Ts is set to t1. Then, MAP2 checks the previous SMAP’s IP address carried in REREQ. If it is null (that is, there is no previous SMAP for MS1), the procedure proceeds to the next step. Otherwise, MAP2 updates the entry for MS1 in its proxy table as follows: Im is set to MS1’s IP Address, Is is set to MAP2’s IP Address, and Ts is set to t1.

Step RE3. MAP2 responds to MAP2 an update response message, URSP. Then, MAP2 sends a registration response message, RERSP, to MS1, which indicates that the registration request has been completed.

Note that MS1 may have ongoing sessions during the movement, handoff management is required to ensure session continuity. Existing handoff management mechanisms such as enhanced Inter-Access Point Protocol (IAPP) [12] may be adopted in this procedure, where packets are buffered in MAP1 and then forwarded to MAP2. This paper concentrates the study on the location management. The details of handoff management are not included in this paper.

After Step RE3, MS1’s location information is kept in the proxy tables of both MAP1 and MAP2, and the location management for MS1 is done at MAP1 and MAP2. For other MNs with obsolete MS1’s location information (that is, Is field for MS1 stores MAP’s IP address), the packets are first routed to MAP1, and then MAP1 retrieves its proxy table to forward the packets to MAP2.

Note that there is no possibility of any loop that resulted from the registration procedure. Suppose that MS1 has the movement, MAP1, MAP2, ..., MAPn, MAP1. When MS1 moves from MAPn to MAP1, the registration procedure is exercised between MS1 and MAP1, which updates MAP1’s proxy table. Since MAP1 is the SMAP of MS1, all packets can be routed to MS1 directly through MAP1. 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 MS1 (sender) is sending IP packets to MS2 (receiver), where MAP3 is one of the MNs along the routing path. Fig. 4a illustrates the flow chart for this part. Steps L1 and L2 update the location information for the sender (that is, MS1).

Step L1. Upon receipt of an IP packet, MAP3 first checks the prefix of MS1’s IP address to determine whether MS1 is in the WMN. If the packet is sent from Internet into the WMN (that is, MS1 is out of the WMN), MAP3 does not need to maintain MS1’s location information, and the procedure jumps to Step L3. Otherwise, the procedure proceeds to the next step.

Step L2. MAP3 checks the options field in the IP header. Two cases are considered.

Case L2.1. The options field is null, that is, MAP3 is MS1’s SMAP, whose proxy table contains MS1’s current location information. MAP3 updates the options field in the IP header: ISS is set to the Is value of MS1’s entry in MAP3’s proxy table; SST is set to the Ts value of MS1’s entry in MAP3’s proxy table; IRS is set to null; RST is set to the null.

Case L2.2. The options field is not null. If MS1’s entry exists in MAP3’s proxy table, MAP3 updates MS1’s location information. Otherwise, an entry is created for MS1 in MAP3’s proxy table. MS1’s entry in MAP3 is set as follows: Im is set to MS1’s IP Address, Is is set to the ISS value in the IP header, and 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 (that is, MS2):

Step L3. MAP3 checks the prefix of MS1’s IP address to determine whether MS2 is in the WMN. If MS2 is out of the WMN, the procedure exits. Otherwise, the procedure proceeds to the next step.
Fig. 4. The flow chart for the routing procedure.
Step 4. This step synchronizes MS’s location information carried in the IP header and that stored in the proxy table. Let \(t_s\) be the TS value in MS’s entry and \(t_p\) be the RST value in the IP header. Without loss of generality, if MS’s entry does not exist, \(t_s = 0\), and if the RST value is null, \(t_p = 0\). We consider three cases:

Case L4.1. \(t_s < t_p\) that is, MS’s location information carried in the IP header is fresher than that stored in the proxy table. MS’s location information in MAP’s proxy table is updated as MS’s location information carried in the IP header. MS’s entry in MAP’s proxy table is updated as follows: Im is set to MS’s IP Address, Is is set to the RSS value in the IP header, and Ts is set to the RST value in the IP header.

Case L4.2. \(t_s = t_p\). MS’s location information carried in the IP header is the same as that in MAP’s proxy table. The procedure does nothing.

Case L4.3. \(t_s > t_p\) that is, MS’s location information in the proxy table of MAP is fresher than that carried in the IP header. MS’s location information carried in the IP header is filled with MS’s location information in MAP’s proxy table. The options field in the IP header is filled as follows: ISS is not changed, SST is not changed, IRS is set to the Is value of MS’s entry in MAP’s proxy table, and RST is set to the Ts value of MS’s entry in MAP’s proxy table.

After Part 1 finishes, the sender’s (that is, MS’s) current location information is stored in MAP’s proxy table and in the IP header. MS’s location information stored in MAP’s proxy table and that carried in the IP header are synchronized.

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 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’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’s SMAP replies a route response message, RRES(IP Address of MS’s SMAP, MS’s Serving time stamp), to the mesh backhaul.

Step Q3. If the RRES message is received before \(T_q\) expires, the mesh backhaul updates MS’s location information carried in the IP header and that in the proxy table. After query procedure, MAP can route the packet. Otherwise (that is, \(T_q\) 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 (that is, the IP address of the receiver’s SMAP) cannot 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 Fig. 5. Suppose that MS0 enters the WMN at \(t_0\). Let \(x\) be the time period between \(t_0\) and the time when MS0 originates the first Internet session. Suppose that the Internet session arrivals originated by MS0 form a Poisson process with rate \(\lambda\). 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 MS0 enters the WMN, there are another \(N\) MSs. We assume that the \(N\) MSs are identical, and each MS initiates intranet sessions toward MS0 with probability \(\gamma\). Let \(N^*(0 \leq N^* \leq N)\) be the number of MSs (that initiate intranet sessions toward MS0). Without loss of generality, we assume that the \(N^*\) MSs are MS1, MS2, …, 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 first intranet session.
The $A$ probability is derived as follows:

$$A = \sum_{j=0}^{N} \Pr[\forall k \in \{1, 2, \ldots, N\}, x \leq y_k | N' = j]$$

$$= \sum_{j=0}^{N} \Pr[\forall k \in \{1, 2, \ldots, j\}, x \leq y_k | N' = j]$$

$$= \sum_{j=0}^{N} \left\{ \int_{x=0}^{\infty} \left[ \prod_{k=1}^{j} f_y(y_k) dy_k \right] f_x(x) dx \right\}$$

$$= \sum_{j=0}^{N} \left\{ N \gamma^j (1-\gamma)^{N-j} \right\}\frac{F_y(x)}{f_y(y_j)} dx$$

$$= \sum_{j=0}^{N} \left\{ N \gamma^j (1-\gamma)^{N-j} \right\}\frac{\lambda e^{-\lambda x} dx}{f_y(y_j)}$$

$$= \sum_{j=0}^{N} \left\{ \lambda \int_{x=0}^{\infty} [1 - F_y(x)]^j e^{-\lambda x} dx \right\}$$

$$= \sum_{j=0}^{N} \left\{ \gamma^j (1-\gamma)^{N-j} \right\}\frac{\lambda^j e^{-\lambda x} dx}{f_y(y_j)}$$

$$= \frac{\lambda}{f_y(y_j)} \int_{x=0}^{\infty} e^{-\lambda x} dx$$

Let $F'_y(y) = [1 - F_y(y)]^j$ and $f'_y(s)$ be the Laplace transform of $F'_y(y)$. Then, (2) is rewritten as

$$A = \lambda \sum_{j=0}^{N} f'_y(\lambda) \gamma^j (1-\gamma)^{N-j}. \tag{3}$$

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

$$\tilde{P}_q = 1 - \lambda \sum_{j=0}^{N} f'_y(\lambda) \gamma^j (1-\gamma)^{N-j}. \tag{4}$$

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

$$F'_y(y) = [1 - (1 - e^{-\eta y})]^j = e^{-j\eta y} \tag{5}$$

and

$$f'_y(s) = \int_{0}^{\infty} e^{-j\eta y} e^{-\eta y} dy = \frac{1}{j\eta + s}. \tag{6}$$

Applying (6) into (4), we have

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

This study also conducts simulation experiments to investigate the $\tilde{P}_q$ 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 [13], [14], [15], [16], [17]. Following the standard [3], a WMN is modeled as a regular hexagonal topology. Each hexagon represents the coverage area of a MAP. In our simulation, the WMN consists of 61 MNs (that is, 1 mesh backhaul + 60 MAPs) and 1,000 MSs (that is, $N = 1,000$). The mesh backhaul is located at the center of the WMN. The movement of an MS follows a 2D random walk model [18], where an MS resides in a MAP’s coverage area for a period and then moves to one of its neighboring
Effects of input parameters are discussed below. The expected MAP residence time is 10 sec. The input parameters are $\xi_k$ and $\eta$. As $\eta$ increases, other MSs initiate their first intranet session to MS0 earlier (that is, $1/\eta$ is smaller). It is more likely that an MS initiates an intranet session to MS0 before MS0 initiates the first Internet session, and the mesh backhaul may not contain MS0’s location information when it routes the intranet session for MS0. Therefore, we observe the larger $P_q$ values when $\eta$ increases.

In the following, we run simulation experiments to investigate the $P_q$ performance for WMM, where we assume $x$ (that is, the time period between $t_0$ and the time when MS0 originates the first Internet session) and $y_k$ (that is, the time period between $t_0$ and the time when MS0, $1 \leq k \leq N'$, originates the first intranet session toward MS0) have exponential distributions with means $1/\lambda$ and $1/\eta$, respectively. In our experiments, we assume that $\eta < \lambda$ (or $1/\eta > 1/\lambda$). 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 MSs that initiate an intranet session toward MS0 in the WMN before it enters WMN. The MS resident time $z$ is assumed to be Gamma distributed with mean $1/\omega$ and variance $v_z$, respectively. The Gamma distribution has widely been adopted to simulate MS moving behavior in the real mobile networks in several studies [19], [13], [20], [21]. The input parameters $\eta$ and $\omega$ are normalized by $\lambda$. For example, if $1/\lambda = 1$ second, then $\omega = 0.1\lambda$ means that the expected MAP residence time is 10 sec. The impacts of the input parameters are discussed below.

- Effects of $\eta$: Fig. 6 plots $P_q$ as an increasing function of $\eta$, where we set $\omega = 0.1\lambda$, $N = 1,000$, and $v_z = 1/\omega^2$ (that is, exponential MAP residence time). As $\eta$ increases, other MSs initiate their first intranet session to MS0 earlier (that is, $1/\eta$ is smaller). It is more likely that an MS initiates an intranet session to MS0 before MS0 initiates the first Internet session, and the mesh backhaul may not contain MS0’s location information when it routes the intranet session for MS0. Therefore, we observe the larger $P_q$ values when $\eta$ increases.

- Effects of $\omega$: Fig. 7 also studies the effects of $\omega$, where $\eta = 0.001\lambda$, $v_z = 1/\omega^2$ (that is, exponential MAP residence time), and $N = 1,000$. As $\omega$ increases, the $P_q$ values slightly drop. With the higher MS mobility, through the registration procedure, the MS’s location information is 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 $\omega$ increases.

- Effects of $v_z$: Fig. 8 studies the effects of $v_z$, where $\eta = 0.001\lambda$, $\omega = 0.1\lambda$, and $N = 1,000$. The figure shows that, with larger $v_z$, the $P_q$ values drop. This is due to the fact that, as $v_z$ becomes large, more small $z$ values are observed. Therefore, the higher MS mobility is expected. Similar to the effects of $\omega$ (see Fig. 7), the MS’s location information is more likely to be cached in the MAPs. We observe that WMM functions better (that is, 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 the ad hoc routing protocol, the centralized-database MM protocol, and the mobile IP protocol) in terms of location update, location tracking, and packet routing cost. More nodes (involved in a location update operation and a location tracking operation) result in more signaling messages replicated within the WMN (that is, more signaling traffic) and a longer routing path for signaling message delivery (that is, the possibility for successful message delivery decreases). Furthermore,
Fig. 8. Effects of variance $v_z (\eta = 10^{-3}; \omega = 0.1; N = 1,000)$.

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 $C_r \leq 3$ 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 the total number of MNs (that is, $M$). 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, $C_q$ for WMM can be estimated as $\frac{M \cdot r}{100}$.

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 $C_q$ 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 $C_q$ is $\hat{R}$.

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 MAP$_i$s, where the routing path between MAP$'_2$ and MAP$_2$ may be a direct link or polygon links. Let $r_1$ be the average number of MNs along the routing path between MAP$_1$ and MAP$'_2$, and $r_2$ be the average number of MNs along the routing path between MAP$'_2$ and MAP$_2$. In this case, we have $C_q = r_1 + r_2 > \hat{R}$.

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, $C_q = r_3 + r_4 > \hat{R}$.
Let the probability that Case 1 occurs be $\beta_1$, the probability that Case 2 occurs be $\beta_2$, and the probability that Case 3 occurs be $\beta_3$. We have

$$\beta_1 + \beta_2 + \beta_3 = 1.$$ (8)

In most of the cases in the real system, communications between two MSs are bidirectional (that is, both MS1 and MS2 exchange packets with each other). Once MS1 and MS2 exchange packets with each other, MS1’s SMAP caches MS2’s current location information, and vice versa, and the prolonged routing path (incurred 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 $R$. Suppose that MS1 sends the first packet at time $t_0$, MS1 and MS2 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

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

Since, in most of applications (for example, TCP session), the bidirectional communication usually starts after MS2 receives the first packet from MS1, we have $t_1 \approx t_0$, that is, $\frac{t_1 - t_0}{t_2 - t_0} \approx 0$ and $\frac{t_2 - t_1}{t_2 - t_0} \approx 1$. With (8), $C_r$ in (9) approximates to $R$.

### 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 [4]. 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. Typically, the MS is close to the FA, and we have $C_o = M + N$ 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 overhead is incurred. In the reactive 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 MS’s movement, and we have $C_o = M + N$. For the reactive ad hoc routing protocol, the first packet at time $t_0$, MS1 and MS2 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

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

Since, in most of applications (for example, TCP session), the bidirectional communication usually starts after MS2 receives the first packet from MS1, we have $t_1 \approx t_0$, that is, $\frac{t_1 - t_0}{t_2 - t_0} \approx 0$ and $\frac{t_2 - t_1}{t_2 - t_0} \approx 1$. With (8), $C_r$ in (9) approximates to $R$.

### 4.3 Signaling and Routing Cost of the Centralized-Database MM Protocol

In the centralized-database MM protocol, a centralized MM database is maintained to store the location information for all MSs. Whenever an MS moves from one 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 $C_o = R$ 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 $C_r = R$ 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 $C_r = R$ for the centralized-database MM protocol.

### 4.4 Signaling and Routing Cost of Mobile IP Protocol

In the 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 $C_o = R$ 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 are no signaling messages required for location tracking, and we have $C_r = 0$ for the mobile IP protocol.

In the mobile IP protocol, the packets are always routed to the HA and then to the destination (that is, the triangle routing problem) and the $C_r$ cost for the mobile IP protocol can be estimated as $2R$.

The previous study [7] proposed route optimization for the mobile IP protocol to overcome triangle routing. Suppose that a corresponding node MS1 is sending packets...
TABLE 1
Comparison between WMM and Other MM Mechanisms

<table>
<thead>
<tr>
<th></th>
<th>The $C_u$ Cost</th>
<th>The $C_t$ Cost</th>
<th>The $C_r$ Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proactive Ad-hoc Routing Protocol</td>
<td>$M + N$</td>
<td>0</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>Reactive Ad-hoc Routing Protocol</td>
<td>0</td>
<td>$\frac{M + N}{n_s}$</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>Centralized-database MM Protocol</td>
<td>$\bar{R}$</td>
<td>$\bar{R}$</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>Base Mobile IP Protocol</td>
<td>$\bar{R}$</td>
<td>0</td>
<td>$2\bar{R}$</td>
</tr>
<tr>
<td>Mobile IP Protocol with Route Optimization</td>
<td>$\bar{R}$</td>
<td>$\bar{R}$</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>WMM Mechanism</td>
<td>$\leq 3$</td>
<td>$\frac{M \cdot P_q}{r}$</td>
<td>$\approx \bar{R}$</td>
</tr>
</tbody>
</table>

5 CONCLUDING REMARKS
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 (that is, fast and low signaling cost) route packets to mobile users. The fields in the 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 cannot 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 the following:

- When an MS enters the WMN, if the first intranet session is initiated toward the MS earlier, the $P_q$ is higher. On the other hand, if the MS initiates its first Internet session earlier, the $P_q$ is lower.
- With higher MS mobility or higher variance of the MAP residence time, the WMM gains better $P_q$ performance (that is, $P_q$ is lower).

At the end of the paper, we made a comparison between WMM and other existing MM protocols (including the ad hoc routing protocol, the centralized-database MM protocol, and the 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.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Chi-Chang Hsieh and Mr. Yen-Ming Chen for their help in implementing the WMM mechanism. They also thank the editor and the three
anonymous reviewers for their valuable comments. Their efforts have significantly improved the quality of this paper. P. Lin’s work was sponsored in part by the National Science Council (NSC), ROC, under contract numbers NSC-96-2627-E-002-001-, NSC-96-2811-E-002-010, NSC-96-2628-E-002-002-MY2, and NSC-95-2221-E-002-091-MY3, the Ministry of Economic Affairs (MOEA), ROC, under contract number 93-EC-17-A-05-S1-0017, the Telcordia Applied Research Center, the Taiwan Network Information Center (TWNIC), the Excellent Research Projects of National Taiwan University, 95R0062-AE00-07, and the Chunghwa telecom M-Taiwan program M-Taoyuan Project. C.-H. Gan’s work was sponsored in part by the NSC, Taiwan, under contract numbers NSC-95-2219-E-009-019- and 95-2218-E-009-201-MY3, the Taiwan Network Information Center (TWNIC), and the Information & Communications Research Labs, Industrial Technology Research Institute (ICL/ITRI).

REFERENCES


Di-Wei Huang received the BS degree in computer science and information engineering from National Chiao-Tung University, Taiwan, in 2004, and the MS degree in computer science and information engineering from National Taiwan University, Taiwan, in 2006. He is currently on active duty of obligatory military service in Taiwan. His current interests include wireless networking and mobile computing.

Phone Lin received the BS degree from National Chiao Tung University, Taiwan, ROC, in 1996 and 2001, respectively. From August 2001 to July 2004, he was an assistant professor in Department of Computer Science and Information Engineering (CSIE), National Taiwan University, ROC. Since August 2004, he has been an associate professor in the Department of CSIE and at the Graduate Institute of Networking and Multimedia, National Taiwan University, ROC. His current research interests include personal communications services, wireless Internet, and performance modeling. Dr. Lin has published more than 20 international SCI journal papers (most of which are IEEE Transactions and ACM papers). He is an associate editor for the IEEE Transactions on Vehicular Technology, a guest editor for the IEEE Wireless Communications special issue on mobility and resource management, and a guest editor for the ACM/Springer MONET special issue on wireless broadband access. He is also an associate editorial member for Wireless Communications and Mobile Computing. Dr. Lin has received many research awards. He was elected as the Best Young Researcher for the Third IEEE ComSoc Asia-Pacific Young Researcher Award, 2007. He was a recipient of the Research Award for Young Researchers from Pan Wen-Yuan Foundation in Taiwan in 2004, the K.T. Li Young Researcher Award honorary by ACM Taipeh Chapter in 2004, the Wu Ta You Memorial Award of the National Science Council (NSC) in Taiwan in 2005, the Fu Suu-Nien Award of NTU in 2005 for his research achievements, and the 2006 Young Electrical Engineering Researcher Award honored by ACM Taipei Chapter in 2006. Dr. Lin is a senior member of the IEEE. For more information, see http://www.csie.ntu.edu.tw/~plin.

Chai-Hien Gan received the BS degree in computer science from Tamkang University, Taipei County, Taiwan, in 1994 and the MS and PhD degrees in computer science and information engineering from National Taiwan University, Taipei, in 1996 and 2005, respectively. From March 2005 to July 2007, he was a research assistant professor in the Department of Computer Science, National Chiao Tung University, Taiwan. Since July 2007, he has been a researcher in the Information and Communications Research Labs, Industrial Technology Research Institute, Taiwan. His current research interests include wireless and mobile computing, personal communications services, IP multimedia subsystem, and wireless Internet. He is a member of the IEEE.