# Dynamic Hierarchical Mobility Management Strategy for Mobile IP Networks

Wenchao Ma, *Member, IEEE,* and Yuguang Fang, *Senior Member, IEEE*

***Abstract*—One of the major challenges for the wireless network design is the efficient mobility management, which can be addressed globally (macromobility) and locally (micromobility). Mobile Internet protocol (IP) is a commonly accepted standard to address global mobility of mobile hosts (MHs). It requires the MHs to register with the home agents (HAs) whenever their care-of addresses change. However, such registrations may cause excessive signaling traffic and long service delay. To solve this problem, the hierarchical mobile IP (HMIP) protocol was proposed to employ the hierarchy of foreign agents (FAs) and the gateway FAs (GFAs) to localize registration operations. However, the system performance is critically affected by the selection of GFAs and their reliability. In this paper, we introduce a novel dynamic hierarchical mobility management strategy for mobile IP networks, in which different hierarchies are dynamically set up for different users and the signaling burden is evenly distributed among the network. To justify the effectiveness of our proposed scheme, we develop an analytical model to evaluate the signaling cost. Our performance analysis shows that the proposed dynamic hierarchical mobility management strategy can significantly reduce the system signaling cost under various scenarios and the system robustness is greatly enhanced. Our analysis also shows that the new scheme can outperform the Internet Engineering Task Force mobile IP hierarchical registration scheme in terms of the overall signaling cost. The more important contribution is the novel analytical approach in evaluating the performance of mobile IP networks.**

***Index Terms*—Mobile IP (MIP), mobility management, roaming, wireless networks.**

## I. Introduction

THE CURRENT fast increasing demand for wireless access to Internet applications is fueled by the remarkable success of wireless communication networks and the explosive growth of the Internet. The future generation wireless networks target to provide users with high-speed Internet access and multimedia services besides voice. The user mobile devices, such as wireless laptops, cellular phones, and palm pilots, make it possible for mobile users to access the Internet applications that are predominantly based on Internet protocol (IP) [1], [2]. The popularity of the Internet provides strong incentive to service providers to support seamless user mobility. However, many telecommunication systems such as first and second generation wireless cellular systems were designed mainly for voice services, and the integration with data networks becomes the major push for third and future generation wireless systems. Mobile IP (MIP) is the mobility-enabling protocol developed by the Internet Engineering Task Force (IETF) to support global mobility in IP networks [3], [4]. This standard has become the solution to solve the user mobility in almost all packet-based wireless mobile systems.

The IP protocol has been designed for wired networks. There are two major functions for the terminal IP addresses in the Internet. An IP address is used to identify a particular end system in the whole network and is also used to find a route between the endpoints. In IP networks, the packets delivered to a particular end system is routed according to the destination IP address by the intermediate routers. Based on this observation, we can conclude that a mobile terminal needs to have a stable IP address in order to be stably identifiable to other network nodes and also needs a temporary IP address for the routing purpose. The MIP protocol extends IP by allowing a mobile node to effectively utilize two IP addresses, one for identification and the other for routing.

MIP enables mobile terminals to maintain ongoing communications with the Internet while moving from one subnet to another. In the MIP protocol, mobile terminals that can change their points of attachment in different subnets are called mobile hosts (MHs). An MH has a permanent address (home address) registered in its home network and this IP address remains unchanged when the user moves from subnet to subnet. This address is used for identification and routing purpose, which is stored in a home agent (HA). An HA is a router in a mobile node's home network, which can intercept and tunnel the packets for the mobile node and also maintains the current location information for the mobile node. If an MH roams to a subnetwork other than the home network, this subnetwork is a foreign network for that user. In the current MIP protocol, the MH can obtain a new IP address from a router [foreign agent (FA)] in the visited network or through some external means. An MH needs to register with the FA or some one-hop router for the routing purpose. The care-of-address (CoA) for the MH will change from subnet to subnet. In order to maintain continuous services while the user is on the move, MIP requires the MH to update its location to its HA whenever it moves to a new subnet so that the HA can intercept the packets delivered to it and tunnel the packets to the user's current point of attachment.

Manuscript received February 1, 2003; revised September 29, 2003. This work was supported in part by the U.S. National Science Foundation under Faculty Early Career Development Award ANI-0093241 and under Grant ANI-0220287, and in part by the U.S. Office of Naval Research under Young Investigator Award N000140210464 and under Grant N000140210554.

W. Ma is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA. He is also with Microsoft Research Asia, Beijing 100080, China (e-mail: mawench@ufl.edu; t-wenma@microsoft.com).

Y. Fang is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: fang@ece.ufl.edu).

Digital Object Identifier 10.1109/JSAC.2004.825968

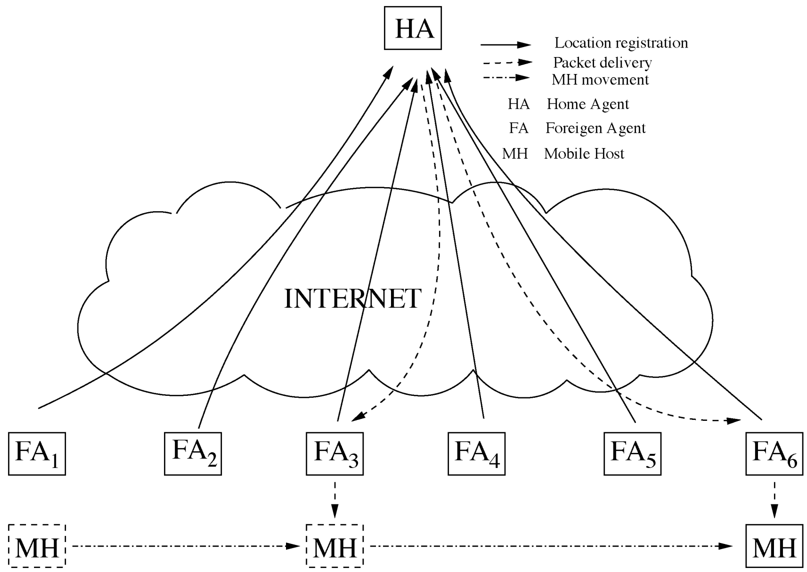



Fig. 1. MIP location registration and packet routing.

Thus, the MIP can provide continuous Internet access services for the mobile user and does provide a simple and scalable solution to user mobility.

However, MIP is not a good solution for users with high mobility. Its mechanism requires every MH to update its new CoA to the HA every time the MH moves from one subnet to another, even though the MH does not communicate with others while moving. As shown in Fig. 1, the location update cost in MIP can be excessive, especially for the mobile users with relatively high mobility and long distance to their HAs. This problem becomes worse with the increase of the mobile user population [5]. Moreover, if a user is far away from his/her HA or the HA processing capability is overwhelmed by the huge volume of update messages, the signaling delay for the location update could be very long, which will result in the loss of a large amount of in-flight packets and the degradation of quality-of-service (QoS).

In this paper, we propose a dynamic hierarchical mobility management scheme for MIP networks (DHMIP). In our scheme, the location update messages to the HAs can be reduced by setting up a hierarchy of FAs, where the level number of the hierarchy is dynamically adjusted based on each mobile user's up-to-date mobility and traffic load information. Analytical model is developed for the performance evaluation. The results show that our new scheme outperforms the MIP and IETF hierarchical mobile IP (HMIP) schemes under various conditions. The more important contribution is the new analytical approach we develop in this paper. Most performance evaluation of mobility management schemes in MIP networks are carried out by simulations, while this work presents a novel analytical approach to the performance evaluation for MIP networks.

This paper is organized as follows. We introduce the related works on the mobility management for the MIP networks in Section II. The detailed procedures of the DHMIP scheme is presented in Section III. In Section IV, we develop an analytical model to derive the signaling cost functions for the new scheme. The performance of DHMIP scheme is demonstrated in Section V. Section VI compares our new scheme with the IETF HMIP scheme and some improvements for the DHMIP are discussed in Section VII. Section VIII gives the conclusions.

## II. Related Works

User mobility in wireless networks that support IP mobility can be broadly classified into macromobility and micromobility. The macromobility is for the case when an MH roams across different administrative domains of geographical regions. The macromobility occurs less frequent and usually involves longer timescales ([6]). The MIP can ensure the mobile users reestablish communication connections after a movement. The micromobility means the MH movement across multiple subnets within a single network of domain. For micromobility, which occurs quite often, the MIP paradigm needs to be enhanced. Most of the related works attempt to improve the MIP micromobility handling capability [7].

In [5], the authors proposed a scalable mobility management scheme which uses hierarchical FAs to handle the user mobility within one subnetwork for wireless Internet, and FA hierarchy in this scheme is preconfigured. In this architecture, the base stations are assumed to be network routers. The higher levels of the hierarchy rely on MIP to handle the macromobility. The handoff-aware wireless access Internet infrastructure (HAWAII) is a separate routing protocol to handle micromobility [8]. The scheme hinges on the assumption that most user mobility is local to an administrative domain of the networks. An MH entering a new foreign network is assigned a new CoA and retains its CoA unchanged while moving within the foreign domain. In this scheme, the HA and any corresponding host are unaware of the host's mobility within that domain. The route to the MH is established by specialized path setup scheme that updates the forwarding tables with host-based entries in selected routes in that domain. For macromobility, the HAWAII uses the traditional MIP. By this means, the scheme can be considered as an enhanced MIP. The cellular IP is introduced in [9]. In this scheme, the location management and handoff support are integrated with routing in cellular IP access networks. The network is connected to the Internet through gateway routers and the roaming between gateways is managed by MIP while the mobility within access networks is handled by cellular IP. All the packets originated from or terminated to the MHs are handled by the gateways, and the host location information is refreshed by the regular data exchange transmitted from the MH. The cellular IP also supports IP paging. The paging mechanism can minimize the signaling traffic for users in the idle state. When packets need to be sent to an idle MH, the host is paged and becomes active. To localize the signaling traffic in supporting IP services in cellular networks, Das *et al.* proposed to use the mobility agent (MA) to reduce the update traffic, leading to a new architecture called TeleMIP [6]. In the TeleMIP, an MA is in charge of one region, handling the CoAs for those MHs roaming in the region. Whenever an MH switches FA in the same region, it only updates the MA. When an MH crosses a region boundary, the MH registers with the new MA, which then sends update message to the HA. It has been demonstrated that TeleMIP does enhance the performance in the IP support over the cellular networks. However, this new architecture may add the network management entity and complexity. In the MIP networks, to reduce the registration signaling traffic (for binding update), the MIP regional registration is proposed in [10], which is called IETF hierarchical scheme in this paper. The protocol employs the FA hierarchy to localize the registration traffic. In this protocol, the HA registers the publicly routable address of the gateway FA (GFA) and the MHs location update messages establish tunnels in a regional network along the path from MHs to GFA. Although multiple levels of hierarchy are mentioned in [10], typically one level architecture, where all FAs are connected to the GFAs, is used for discussions. In the IETF hierarchical scheme, the network architecture is centralized. It is not clear and usually hard to determine the size of a regional network. The mobile users up-to-date traffic load and mobility may vary, and the fixed structure is lack of flexibility. To overcome this deficiency, Xie and Akyildiz proposed a distributed dynamic regional location management scheme for MIP [11]. In this scheme, the first FA an MH registers at in a new regional network is selected as the GFA, and the regional network size is adjusted based on the user's current traffic load and mobility information. It can be considered as the extension of the IETF regional registration scheme to make it more flexible and adaptive. Some schemes based on "pointer forwarding" technique are proposed in [19] and [20]. In these schemes, pointers are setup when MHs move to new subnets. However, in [19], the authors assume there is correlation between the communication and the geographic distances and an MHs coordinate need to be known in the mobility management. In this paper, we propose a new scheme which has no restriction on the shape or geographic location of subnets. The pointer chain length is critical to the performance of the pointer forwarding scheme and the threshold should change according to the user's up-to-date mobility and traffic load. However, the impact of the chain length was not considered in [20].

In this paper, we propose another dynamic hierarchical mobility management scheme for the MIP networks. In our scheme, when an MH changes its subnet and obtains a new CoA from the new FA (one-hop router), the new FA (one-hop router) updates the new address to the MHs previous FA (one-hop router) so that the new FA (one-hop router) forms a new location management hierarchical level for that user. The FA hierarchical architecture is specific for every user, which makes the user avoid frequently updating his/her home network. The packets delivered to the MH can be tunneled via the multiple levels of FAs to the user. In order to avoid long packet delivery delay, there is an optimal level number (or threshold) for the hierarchy for each user according to his/her call-to-mobility ratio (CMR). The threshold can be dynamically adjusted based on the up-to-date mobility and traffic load for each terminal. When the threshold is reached, the MH performs home registration and sets up a new hierarchy for its further movements. The optimal threshold for each user can be derived by an iterative algorithm. Since our scheme significantly reduces the registration traffic to the HAs, it can also greatly reduce the in-flight packet loss.

## III. Dynamic Hierarchical Location Management Scheme

In our new dynamic hierarchical system architecture, there is no fixed hierarchical architecture for users or any restriction on the shape and the geographic location for subnets. In the MIP protocol, an MH can determine if it enters a new subnet by detecting the agent advertisement messages sent by the mobility agents (HAs or FAs). The MH then obtains a new CoA and sends a location update message to its HA. Upon receiving the message, the HA sets up a binding between the MH permanent address and current CoA so that the HA can intercept the packets to this MH and tunnel them to the user's current access point. The MHs in the MIP networks are required to update their new CoA whenever they are changing their locations (subnets) even though the MHs do not communicate with others. As shown in Fig. 1, this could result in significant signaling traffic to the networks.

In our DHMIP, the location update signaling traffic can be reduced by registering the new CoA to the previous FA, as shown in Fig. 2. By this procedure, a dynamic location hierarchy is constructed for a specific mobile user. The packets for this user can be intercepted and retunneled along the FA hierarchy to the mobile terminal. Thus, the location update traffic can be localized. In this scheme, we can use the similar procedures in [12] and [13] to notify the previous FA the user's new CoA. However, the forwarding through multiple FAs will cause some service delivery delay, which may not be appropriate when there is delay restraint for some Internet applications such as video or voice services. In order to avoid excessive packet transmission delay, we set a threshold on the level number of the hierarchy in the DHMIP scheme. When the threshold is reached, the MH will register to its HA. In the DHMIP scheme, the threshold is dynamically adjusted based on every user's traffic load and mobility. The operation of the DHMIP scheme can be seen in Fig. 2. In Fig. 2, an MH moves from subnet_1 to subnet_6. In the figure, we assume the threshold level of the hierarchy is three. When the user is in subnet_2, subnet_3, subnet_5, or subnet_6, the

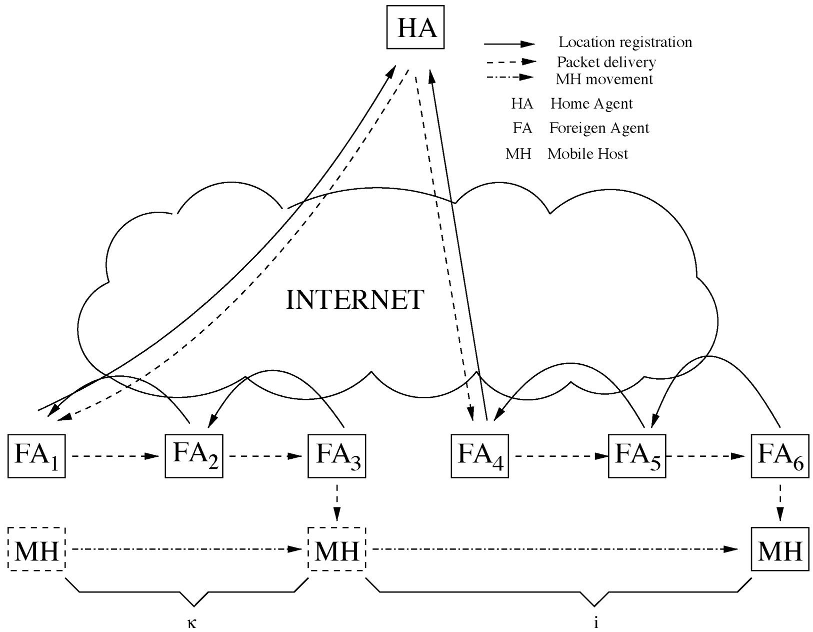



Fig. 2. DHMIP location registration and call delivery.

MH updates the new CoAs to the previous FAs. Since the previous FAs are usually close to the new ones, the location update cost is usually less than that to the HA. After the user enters subnet_4, the hierarchy level threshold is reached and the MH will set up a new hierarchy. In this case, the MH updates its new CoA to the HA directly. When the user is in subnet_3 and subnet_6, there are packet arrivals for the user. The packets are then intercepted by the HA and tunneled to the user. Since the HA does not have the user's up-to-date location information, the packets are sent to the FA that the user updated last time. In our example, they are $\mathrm{FA}_1$ and $\mathrm{FA}_4$ in Fig. 2. Then, the packets are retunneled along the hierarchy to the user. The optimal hierarchy level threshold $K_{\mathrm{opt}}$ can be computed based on the user's current traffic load and mobility pattern. The $K_{\mathrm{opt}}$ can be adjusted in different epochs in the DHMIP scheme. For example, the optimal value can be updated every time the MH enters a new subnet or when the previously calculated threshold is reached or the MH calculates it periodically. There is a tradeoff between the accuracy of $K_{\mathrm{opt}}$ and the MHs computational power consumption. The more often the update of $K_{\mathrm{opt}}$, the more accurate the value and the more the signaling traffic saving; however, the more the power consumption. The DHMIP scheme can be described by the pseudocode in Table I. In this paper, we assume the optimal value is updated every time the last optimal threshold is reached.

TABLE I
DYNAMIC HIERARCHICAL MIP PROTOCOL

```
% Location registration procedures
K_opt: the optimal hierarchy level threshold;
Initialize i = 0;
IF (MH enters a new subnet)
   i = i + 1;
   IF (i ≤ K_opt)
      New FA registers to the previous FA;
   ELSE
      New FA registers to the HA;
      i = 0;
      Compute the new K_opt;
   ENDIF
ENDIF
% Packet delivery procedures
IF (packets for the MH are intercepted by the HA)
   Tunnel the packets to the first FA;
   IF (The first FA is not the MH current serving FA)
      Retunnel the packets to the current FA;
   ENDIF
   The current FA decapsulates the packets and sends them to the MH;
ENDIF
```

## IV. Analytical Model

In this section, we develop an analytical model to derive the location update and packet delivery cost functions for the DHMIP scheme. Just as in [11], we do not consider the periodic binding update costs that an MH sends to its HA or FAs to refresh their caches.

We define the following parameters for our analysis in the rest of this paper.

$\kappa$ Number of subnets crossed between the last packet arrival and the very last location update just before the last packet arrival (see Fig. 2).

$\rho$ MH call-to-mobility ratio (CMR).
$K$ Threshold of the FA hierarchy level.
$U$ Average MH location update cost to its HA.
$F$ Packet delivery cost for a packet to an MH in a foreign network under the MIP scheme.
$M'$ Total location update cost for an MH incurred between two consecutive packet arrivals under the DHMIP scheme.
$F'$ Average packet delivery cost for a packet to an MH in a foreign networks under the DHMIP scheme.
$T$ Packet delivery cost between FAs in the DHMIP scheme.
$S$ Hierarchy setup cost in the DHMIP scheme.
$\beta_\kappa$ Probability distribution of $\kappa$.

Let $\alpha(i)$ denote the probability that an MH crosses $i$ subnets between two consecutive packet arrivals. Under the current MIP scheme, the total location update and packet delivery cost can be shown as

$$C(\rho) = \sum_{i=0}^{\infty} iU\alpha(i) + F = \frac{U}{\rho} + F \tag{1}$$

where the $\rho$ is the MH CMR. In this paper, the CMR is defined as follow: if packets arrive at an MH at rate $\lambda$ and the time the user resides in a given subnet has a mean $1/\mu$, then the CMR, denoted by $\rho$, is given as

$$\rho = \frac{\lambda}{\mu}. \tag{2}$$

The cost for the DHMIP is more complicated, so we derive the location update and packet delivery costs, respectively. By the new scheme, if we assume the threshold is $K$, as we can see in Fig. 2, and if an MH crosses $i$ subnets between two consecutive packet arrivals, the MH will update to the HA $\lfloor (i+\kappa)/K \rfloor$ times and update to the previous FA in the rest $i - \lfloor (i+\kappa)/K \rfloor$ times [14]. So the average location update cost function can be written as

$$M'(\kappa, K, \rho) = \sum_{i=0}^{\infty} \left( \left\lfloor \frac{i+\kappa}{K} \right\rfloor U + \left( i - \left\lfloor \frac{i+\kappa}{K} \right\rfloor \right) S \right) \alpha(i). \tag{3}$$

Similarly, we can obtain the packet delivery cost function. In the DHMIP scheme, some additional cost is introduced. When a packet is tunneled to the first FA in the MH current FA hierarchy, if the first FA is not the user's serving FA, then additional FAs have to be traversed before the packet can reach the destination. In addition to the packet delivery cost in the MIP scheme, there is another $(i+\kappa - \lfloor (i+\kappa)/K \rfloor K)T$ cost in the DHMIP scheme. So the $F'$ can be obtained as follows:

$$F'(\kappa, K, \rho) = F + \sum_{i=0}^{\infty} \left( i + \kappa - \left\lfloor \frac{i+\kappa}{K} \right\rfloor K \right) T\alpha(i). \tag{4}$$

In order to analyze the performance of the new scheme, we need to evaluate $\alpha(i)$ in more detail. We make the following assumptions.

1) The packet arrivals to an MH form a Poisson process with arrival rate $\lambda$.
2) The residence time of an MH in a subnet is a random variable with a general density function $f_m(t)$ and the Laplace transform

$$f_m^*(s) = \int_{t=0}^{\infty} f_m(t) e^{-st} dt.$$

For convenience, we denote $g = f_m^*(\lambda)$. With the above assumption, $\alpha(i)$ is given by ([15])

$$\alpha(i) = \begin{cases} 1 - \frac{1-g}{\rho}, & \text{if } i = 0 \\ \frac{(1-g)^2 g^{i-1}}{\rho}, & \text{if } i > 0 \end{cases}. \tag{5}$$

In order to analyze (3) and (4), we use the substitution $i = jK + q$, then

$$\alpha(jK+q) = \frac{(1-g)^2}{\rho g} (g^K)^j g^q = yz^j x^q \tag{6}$$

where

$$y = \frac{(1-g)^2}{\rho g}, \quad z = g^K, \quad x = g.$$

Notice that both $0 \le q < K$ and $0 \le \kappa < K$, we can rewrite $M'$ as

$$\begin{aligned} M'(\kappa, K, \rho) &= S \sum_{i=0}^{\infty} i\alpha(i) + (U - S) \sum_{i=0}^{\infty} \left\lfloor \frac{i+\kappa}{K} \right\rfloor \alpha(i) \\ &= \frac{S}{\rho} + (U-S) \\ &\quad \times \sum_{j=0}^{\infty} \sum_{q=0}^{K-1} \left\lfloor \frac{jK+q+\kappa}{K} \right\rfloor \alpha(jK+q) \\ &= \frac{S}{\rho} + (U-S) \\ &\quad \times \sum_{j=0}^{\infty} \sum_{q=0}^{K-1} \left\lfloor \frac{jK+q+\kappa}{K} \right\rfloor yz^j x^q \\ &= \frac{S}{\rho} + (U-S)y \\ &\quad \times \sum_{j=0}^{\infty} \left[ jz^j \sum_{q=0}^{K-\kappa-1} x^q + (j+1) z^j \sum_{q=K-\kappa}^{K-1+\kappa} x^q \right] \\ &= \frac{S}{\rho} + \frac{(U-S)y(1-x^{K+\kappa})}{1-x} \\ &\quad \times \sum_{j=0}^{\infty} jz^j + \frac{(U-S)y(x^{K-\kappa} - x^{K+\kappa})}{1-x} \sum_{j=0}^{\infty} z^j \\ &= \frac{S}{\rho} + \frac{(U-S)y(1-x^{K+\kappa})z}{(1-x)(1-z)^2} \\ &\quad + \frac{(U-S)y(x^{K-\kappa} - x^{K+\kappa})}{(1-x)(1-z)} \\ &= \frac{S}{\rho} + \frac{(U-S)(1-g)g^{K-1}}{\rho(1-g^K)} \\ &\quad \times \left[ \frac{1-g^{K+\kappa}}{1-g^K} + \frac{1-g^{2\kappa}}{g^\kappa} \right]. \end{aligned} \tag{7}$$

Similarly, we can obtain

$$\begin{aligned}F'(\kappa,K,\rho) &= F+\sum_{i=0}^{\infty} iT\alpha(i)+\sum_{i=0}^{\infty}\kappa T\alpha(i)\\ &\quad - KT\sum_{i=0}^{\infty}\left\lfloor\frac{i+\kappa}{K}\right\rfloor\alpha(i)\\ &= F+\frac{T}{\rho}+\kappa T-\frac{KT(1-g)g^{K-1}}{\rho(1-g^K)}\\ &\quad\times\left[\frac{1-g^{K+\kappa}}{1-g^K}+\frac{1-g^{2\kappa}}{g^\kappa}\right].\end{aligned} \tag{8}$$

Thus, the total cost is

$$\begin{aligned}C'(\kappa,K,\rho) &= M'(\kappa,K,\rho)+F'(\kappa,K,\rho)\\ &= F+\frac{S+T}{\rho}+\kappa T\\ &\quad+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)}\\ &\quad\times\left[\frac{1-g^{K+\kappa}}{1-g^K}+\frac{1-g^{2\kappa}}{g^\kappa}\right].\end{aligned} \tag{9}$$

Furthermore, if we assume that the $\kappa$ is uniformly distributed with $\beta_\kappa = 1/K$, we can rewrite (9) as

$$\begin{aligned}C'_{\text{uniform}}(K,\rho) &= \frac{1}{K}\sum_{\kappa=0}^{K-1} C'_\kappa\\ &= F+\frac{S+T}{\rho}+\frac{(K-1)T}{2}\\ &\quad+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)K}\\ &\quad\times\sum_{\kappa=0}^{K-1}\left[\frac{1-g^{K+\kappa}}{1-g^K}+\frac{1-g^{2\kappa}}{g^\kappa}\right]\\ &= F+\frac{S+T}{\rho}+\frac{(K-1)T}{2}\\ &\quad+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)K}\\ &\quad\times\left[\frac{K}{1-g^K}+\frac{g^{1-K}-g-1}{1-g}\right].\end{aligned} \tag{10}$$

For demonstration purposes, we assume that the subnet residence time is Gamma distributed with mean $1/\mu$. The reason that Gamma distribution is selected is its flexibility in setting various parameters, it is known that the Gamma distribution can be used to fit the first two moments of the field data. The Laplace transform of a Gamma distribution is

$$f_m^*(s) = \left(\frac{\gamma\mu}{s+\gamma\mu}\right)^\gamma$$

thus, we have

$$g = f_m^*(\lambda) = \left(\frac{\gamma\mu}{\lambda+\gamma\mu}\right)^\gamma = \left(\frac{\gamma}{p+\gamma}\right)^\gamma. \tag{11}$$

In particular, when $\gamma = 1$, we have an exponential distribution for the subnet residence time. If the residence time is exponentially distributed, we have $g = 1/(1+\rho)$, then (10) is reduced to

$$\begin{aligned}C'_{\text{uniform}}(K,\rho) &= F+\frac{S+T}{\rho}+\frac{(K-1)T}{2}\\ &+\frac{U-S-KT}{(1+\rho)^K-1}\left[\frac{(1+\rho)^K}{(1+\rho)^K-1}+\frac{(1+\rho)^K-2-\rho}{K\rho}\right].\end{aligned} \tag{12}$$

In reality, many mobile users usually keep accessing Internet in a vicinity. When they roam far away from their daily work place, they would access the network less often. We can analyze this situations by letting the $\kappa$ linearly or exponentially distributed. If $\kappa$ is linearly distributed with $\beta_\kappa = 2(K-\kappa)/K(K+1)$, then

$$\begin{aligned}C'_{\text{linear}}(K,\rho) &= \sum_{\kappa=0}^{K-1}\beta_\kappa C'_\kappa\\ &= F+\frac{S+T}{\rho}+\frac{(K-1)T}{3}\\ &\quad+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)}\\ &\quad\times\sum_{\kappa=0}^{K-1}\frac{2(K-\kappa)}{K(K+1)}\\ &\quad\times\left[\frac{1-g^{K+\kappa}}{1-g^K}+\frac{1-g^{2\kappa}}{g^\kappa}\right].\end{aligned} \tag{13}$$

If $\kappa$ is exponentially distributed with $\beta_\kappa = (e^{-\kappa}(1-e^{-1}))/(1-e^{-K})$

$$\begin{aligned}&C'_{\text{exponential}}(K,\rho)\\ &= \sum_{\kappa=0}^{K-1}\beta_\kappa C'_\kappa\\ &= F+\frac{S+T}{\rho}+T\left[\frac{e^{-1}+(K-1)e^{-K-1}-Ke^{-K}}{(1-e^{-K})(1-e^{-1})}\right]\\ &\quad+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)}\\ &\quad\times\sum_{\kappa=0}^{K-1}\frac{e^{-\kappa}(1-e^{-1})}{1-e^{-K}}\left[\frac{1-g^{K+\kappa}}{1-g^K}+\frac{1-g^{2\kappa}}{g^\kappa}\right].\end{aligned} \tag{14}$$

For MHs with different traffic load and mobility patterns, their optimal hierarchy level thresholds should be different. The optimal threshold $(K_{\text{opt}})$ for an MH is the value of $K$ that minimizes the cost functions derived above. Since the $K_{\text{opt}}$ must be an integer, we use a similar method proposed in [11] and [17] to obtain the optimal values. We define the cost difference equation between the system with level $K$ and the one with level $K-1 (K \geq 2)$ as

$$\Delta(K,\rho) = C'(K,\rho) - C'(K-1,\rho). \tag{15}$$

Given the CMR, the algorithm to find the optimal value of $K$ is defined as

$$K_{\text{opt}} = \begin{cases}1, & \text{if } \Delta(2,\rho) > 0\\ \max\{K : \Delta(K,\rho)\leq 0\}, & \text{otherwise}\end{cases}. \tag{16}$$

Notice that the algorithm to find the optimal value is iterative. It is easy to implement; however, it may result in local minima. How to avoid the local minima was discussed in [18]. In practical systems, we can determine the optimal value to a limited predefined maximum number. Then, the optimal value can be found by evaluating the total costs for each of the allowed numbers in the range. The estimation of the MH packet arrival rate was also discussed in [17].

## V. Numerical Results

To evaluate the performance of the DHMIP scheme, we need to know the relative cost of updating HA and the previous FA. We use the following notations for our analysis.

- $m_{mf\_u}$ Transmission cost of location update between an MH and an FA.
- $m_{ff\_u}$ Transmission cost of location update between FAs.
- $m_{fh\_u}$ Transmission cost of location update between an FA and the HA.
- $p_{fh\_u}$ Location update processing cost between an FA and the HA.
- $p_{mf\_u}$ Location update processing cost between an MA and an FA.
- $p_{ff\_u}$ Location update processing cost between FAs.
- $m_{ff\_d}$ Transmission cost of packet delivery between FAs.
- $m_{hf\_d}$ Transmission cost of packet delivery between the HA and an FA.
- $m_{fm\_d}$ Transmission cost of packet delivery between an FA and an MH.
- $p_{ff\_d}$ Packet delivery processing cost between FAs.
- $p_{hf\_d}$ Packet delivery processing cost between the HA and an FA.
- $p_{fm\_d}$ Packet delivery processing cost between an FA and an MH.

According to the above definitions and the protocols of the MIP and the DHMIP scheme described in Section III, the MH registration cost $(U)$ to the HA, the packet delivery cost $(F)$ in the MIP protocol, the hierarchy setup cost $(S)$, and the hierarchical packet delivery cost $(T)$ in the DHMIP scheme can be expressed as

$$U = 2m_{mf\_u} + 2m_{fh\_u} + 2p_{mf\_u} + p_{fh\_u} \quad (17)$$

$$S = 2m_{mf\_u} + 2m_{ff\_u} + 2p_{mf\_u} + p_{ff\_u} \quad (18)$$

$$T = m_{ff\_d} + p_{ff\_d} \quad (19)$$

$$F = m_{hf\_d} + m_{fm\_d} + p_{hf\_d} + p_{fm\_d}. \quad (20)$$

The MIP scheme may result in heavy system traffic load because of the MHs' location update messages. The DHMIP scheme attempts to reduce the registrations to the HAs by sending MHs' new location to the old FAs. This would effectively reduce the long distance signaling traffic at the expense of increasing the local traffic load. Usually, an MHs new FA is close to its old one in term of the hops between them and the traffic load is distributed evenly among the system, so that the network can accept more service requests. It is obvious that the performance of DHMIP scheme depends on the relative costs of the long distance registrations to the HA and the local registrations to FAs. In this paper, we use the parameters in Tables II and III for performance analysis. We use three sets of parameters in our analysis. The signaling costs between the adjacent hierarchy levels are fixed and the costs between the HA and an FA are 5, 10, and 20 times of those in the three sets of parameters, respectively.

TABLE II
Performance Analysis Parameters

| $p_{fm\_d}$ | $p_{ff\_d}$ | $p_{hf\_d}$ | $m_{fm\_d}$ | $m_{ff\_d}$ |
|---|---|---|---|---|
| 1 | 0.5 | 10 | 5 | 25 |
| $p_{mf\_u}$ | $p_{ff\_u}$ | $p_{fh\_u}$ | $m_{mf\_u}$ | $m_{ff\_u}$ |
| 2 | 1 | 20 | 10 | 50 |

TABLE III
Performance Analysis Parameters

| set | $m_{fh\_u}$ | $m_{hf\_d}$ |
|---|---|---|
| 1 | 250 | 125 |
| 2 | 500 | 250 |
| 3 | 1000 | 500 |

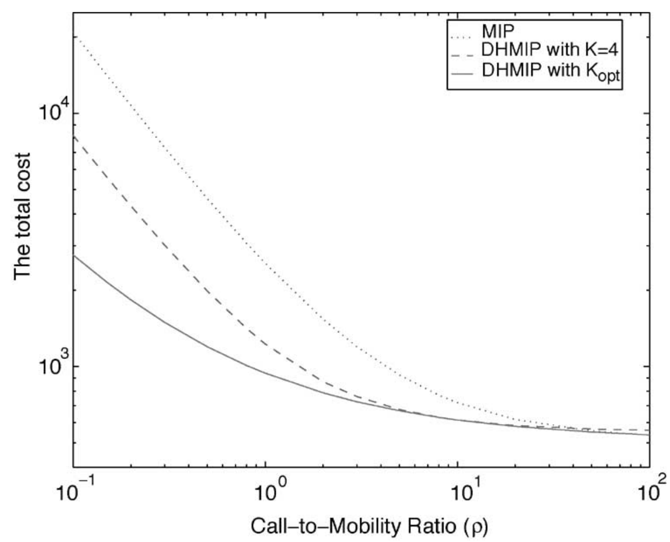



Fig. 3. Comparison of the total costs for different schemes.

Fig. 3 shows the total signaling costs for MIP and the DHMIP scheme under various CMR values. In this figure, we observe that when the CMR is small, the DHMIP scheme generates less traffic than the MIP scheme does even without using the optimal values. We also see that our scheme can reduce the traffic load even further with the optimal values. How to obtain the optimal value has been discussed in Section IV. If we examine Fig. 3 more carefully, we can see that the total cost for the DHMIP scheme with fixed threshold exceeds that for the MIP scheme when the CMR is large. This can be seen more clearly in Fig. 4. In Fig. 4, we show the relative costs of the DHMIP scheme to that of the MIP scheme under different CMR. The uniform distribution is used and the thresholds are four for all the curves in this figure. It is obvious that the total cost for the DHMIP scheme with fixed threshold can exceed that for the MIP scheme when the CMR is large. In the new scheme, the packets for an MH will be processed and tunneled through more FAs before it reaches the destination. When the packet arrival rate is high relative to the mobility and the threshold is fixed, the total cost

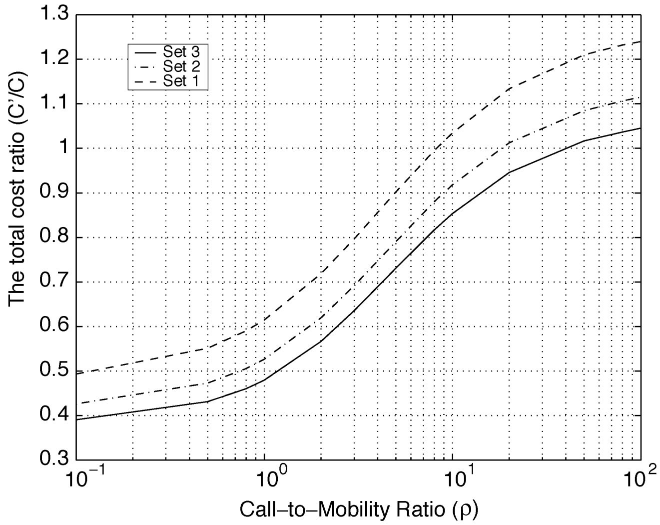



Fig. 4. Comparison of the total costs for uniformly distributed $\kappa$ under different parameters.

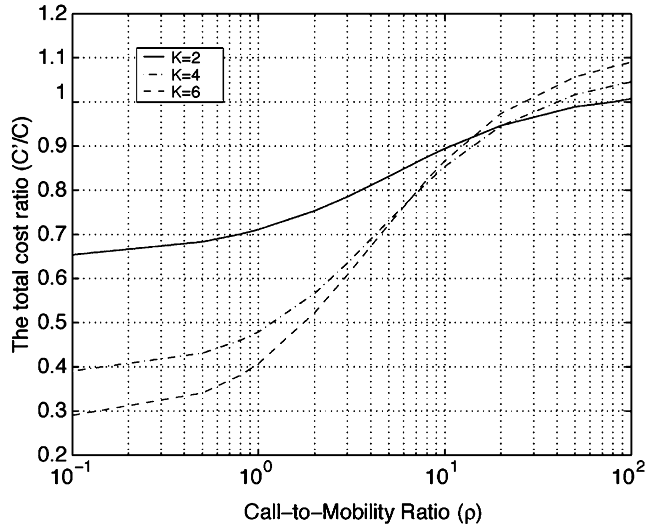



Fig. 6. Comparison of the total costs under different $K$.

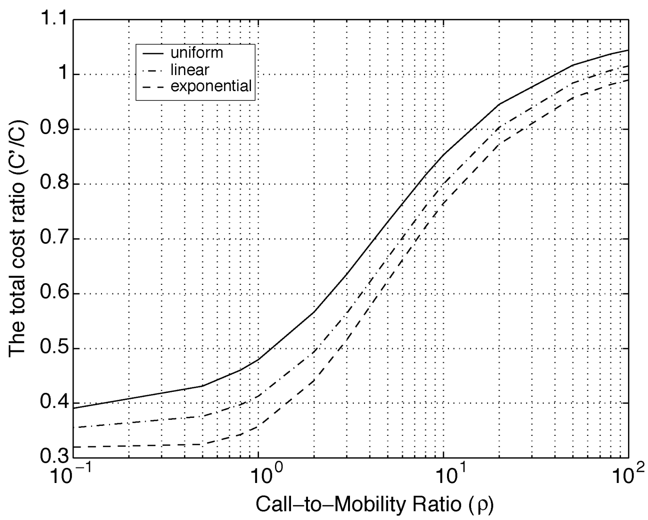



Fig. 5. Comparison of the total costs under different $\kappa$ distributions.

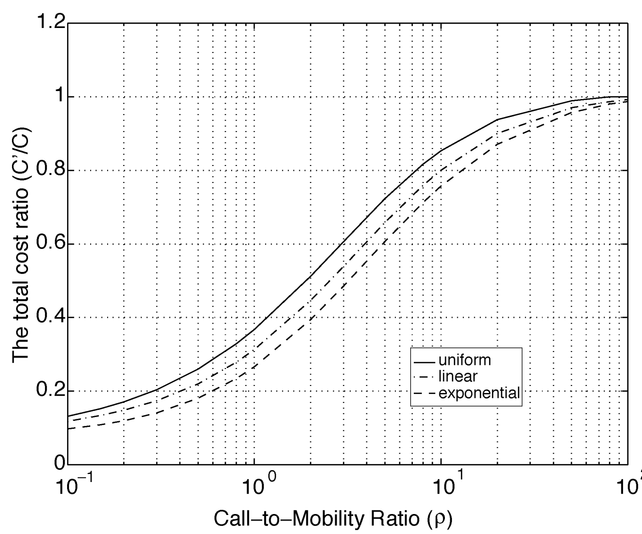



Fig. 7. Comparison of the total costs under different $\kappa$ distributions with optimal $K$.

may exceed that for the MIP scheme. However, we can show that, with the optimal values, the DHMIP scheme will never generate more traffic than the MIP scheme does under any conditions. We can also see in Fig. 4 that the scheme performs better when the relative cost to the HA is higher, and the reason is straightforward.

Some mobile users usually roam only in a vicinity when they access the Internet applications. For example, a passenger waiting for his/her flight in a terminal or the people in some construction field. They usually require less service when they move away from some specific areas. In this paper, we use the linear and exponential distributions to simulate those kinds of mobile users and show the performance in Fig. 5. It can be seen that, with fixed threshold, the users roaming around a specific area can generate less signaling traffic under the DHMIP scheme. In fact, when a mobile user roams in a specific area, there is high probability that the user will revisit some subnets. This will generate loops in the hierarchy and result in signaling traffic which can be removed by "loop removal" introduced in Section VII.

Before we discuss the choice of the optimal threshold for a mobile user, we study the impact of the hierarchy level number on the performance of the DHMIP scheme. Under various CMR values, we change the hierarchy level number $K$ in Fig. 6. When the CMR is low, meaning that the user has relatively high mobility, larger $K$ can minimize the total signaling costs because more updates to the HA can be replaced by hierarchy setup, which will cost less. However, when the CMR is high, the packets delivered to an MH have to travel more FAs before reaching the destination with larger $K$. In this case, the total costs of the DHMIP scheme may exceed those of the MIP scheme. We can also see in Fig. 6 that it is hard to find an optimal fixed threshold for an MH under different situations. In order to achieve the best performance, the selection of threshold must be dynamic.

In Fig. 7, we show the relative performance of the DHMIP scheme using the optimal values under different CMR and different $\kappa$ distributions. We can observe that our new scheme can reduce 85% of the signaling cost when the CMR is small. With

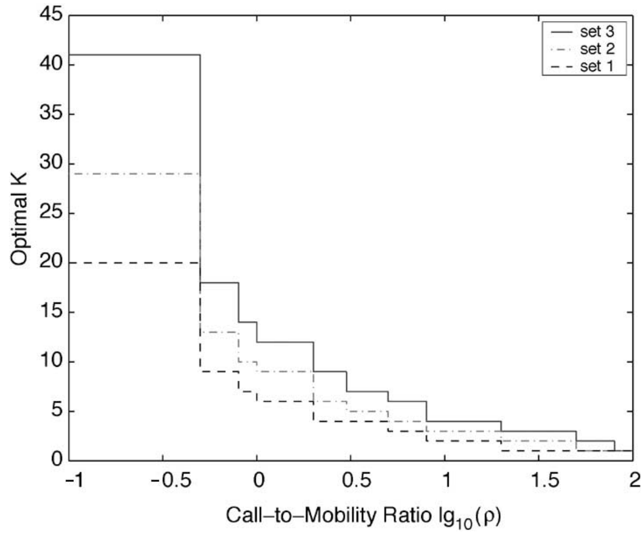



Fig. 8. Optimal number of hierarchy levels for uniform distribution.

the increase of CMR, the relative cost also increases. However, the cost for the DHMIP scheme will never exceed that for the MIP scheme. Fig. 8 shows the optimal values under different CMR for the three sets of parameters. In this figure, we assume the $\kappa$ for the MH follows a uniform distribution. In Fig. 8, the optimal values decrease with the increase of CMR. When the relative update cost to the HA is large, an MH can have relatively large thresholds. The reason is obvious. If the update cost to the HA is high, an MH can generate less signaling traffic by setting up more FA hierarchy, otherwise, the MH should update the HA more often, which implies smaller $K$. In fact, when the optimal value equals one, our new scheme becomes the MIP scheme, so that the total costs will never exceed those for the MIP scheme. Based on the above analysis, we can conclude that the MIP scheme is suitable for mobile users with high CMR and our DHMIP scheme can perform well for all kinds of users.

Next, we investigate the sensitivity of the performance costs and the benefits of the DHMIP scheme to the variance in the mobile user's mobility pattern. In our analysis, we assume that the packet arrivals to a mobile user form a Poisson process and the user's residence time in a subnet has a Gamma distribution. For a Gamma distribution, the variance is $V = \mu^2/\gamma$. That is, a large $\gamma$ implies a small variance. Fig. 9 demonstrates the effects of the variance of the FA residence time under three distributions. In this figure, we assume $\gamma = 0.01$, 1, and 100, respectively. We can see that the variance of the FA residence time in the subnets does not affect the performance of the DHMIP scheme very much. We derive the optimal values based on the exponential residence time distribution. When the $\gamma = 0.01$ or 100, those values are not optimal anymore. However, the DHMIP can still achieve similar performance under different CMR, different residence time and $\kappa$ distributions. This phenomenon proves the flexibility and the adaptivity of our new scheme.

## VI. Comparison With the IETF Hierarchical Scheme

The IETF HMIP protocol aims to reduce the signaling traffic to the home networks while minimizing the signaling delay. The specification of the HMIP protocol can be found in [10]. In this protocol, the HMIP employs the hierarchy of FAs to locally handle MIP registration. When an MH moves from one subnet to another, it performs a home registration to its HA. During a home registration, the HA registers the CoA of the MH. When the visited domain supports regional tunnel management, the CoA that is registered at the HA is the publicly routable address of a GFA. This CoA will not change when the mobile node changes FA under the same GFA. In this situation, the MH will perform a regional update to GFA, the MH mobility management can be handled locally in this way. During the communication, when packets are sent to the MH, the user's HA intercepts the packets, encapsulates them and tunnels them to the CoA of the MH. Those packets will reach the GFA of the MH through the network. The GFA checks its visitor list and retunnels the packets to the user's corresponding FA. The FA further relays the packets to the MH. Typically, one level of hierarchy, where all FAs are connected to the GFAs, is considered; however, the protocol may be utilized to support multiple hierarchy levels, as discussed in Appendix B in [10].

In this section, we compare our DHMIP scheme with the IETF HMIP scheme. We assume $U_1$ is the MH location update cost to a GFA, $U_2$ is that to the HA and $F_{gf}$ is the additional packet delivery cost introduced by the GFA. In the IETF HMIP scheme, the system structure is fixed. If an MH roams under the same GFA, the user total mobility management costs can be reduced. However, the HMIP protocol requires an MH to perform home registration when the user enters a subnet charged by another GFA. In order to analyze the HMIP performance, we need to know the probability that a user moves out of a regional network. We further assume $\omega$ is the probability that the next movement of MH is under the same GFA. In order to obtain $\omega$, we assume there are $N$ subnets in a foreign network and $m$ subnets in one regional area, respectively. In the model, we define the action that each MH moving out of a subnet *a movement*. In one movement, an MH can visit a subnet randomly. Then, we can have $\omega = (m-1)/(N-1)$. Usually, the distance between a GFA and the MHs current FA is farther than the FAs in our DHMIP scheme in term of hops, so that the location update and packet delivery costs are also higher. Conservatively, we define the cost coefficient $\zeta = 2$. The GFA processing cost is proportional to the number of subnets under it. Since IP routing table lookup is based on the *longest prefix matching*, then the GFA processing complexity is proportional to the logarithm of $m$. According to the above assumptions, we define

$$U_1 = 2m_{mf\_u} + 2\zeta m_{ff\_u} + 2p_{mf\_u} + log(m)p_{ff\_u} \quad (21)$$

$$U_2 = 2m_{mf\_u} + 2\zeta m_{ff\_u} + 2m_{fh\_u} + 2p_{mf\_u} + 2log(m)p_{ff\_u} + p_{fh\_u} \quad (22)$$

$$F_{gf} = \zeta m_{ff\_d} + log(m)p_{ff\_d}. \quad (23)$$

In Fig. 10, we plot the curves for the relative cost of the IETF HMIP scheme and our DHMIP scheme to that of the basic MIP protocol. In this figure, we assume $N = 50$, $m = 10$ and 20, respectively. We can see in this figure that our DHMIP scheme generates lower system signaling cost than the IETF HMIP scheme under both conditions even without using the optimal threshold values. We can also see that if the number

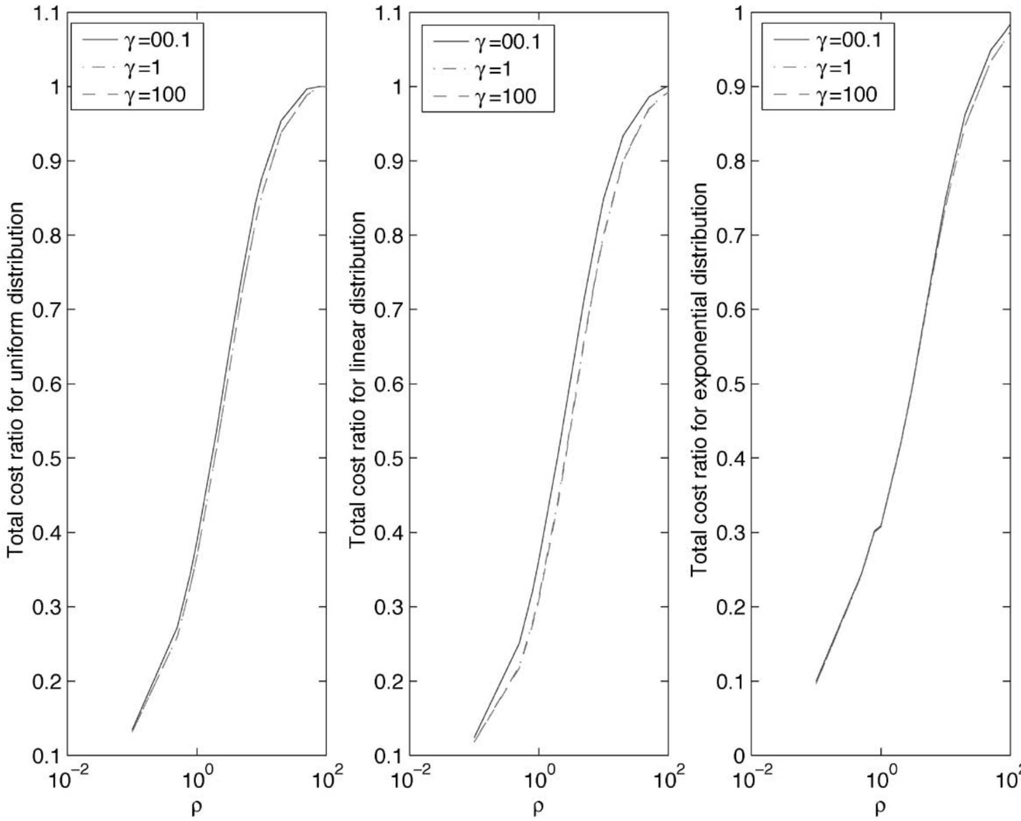



Fig. 9. Effect of the variance of the FA residence time $(\gamma)$.

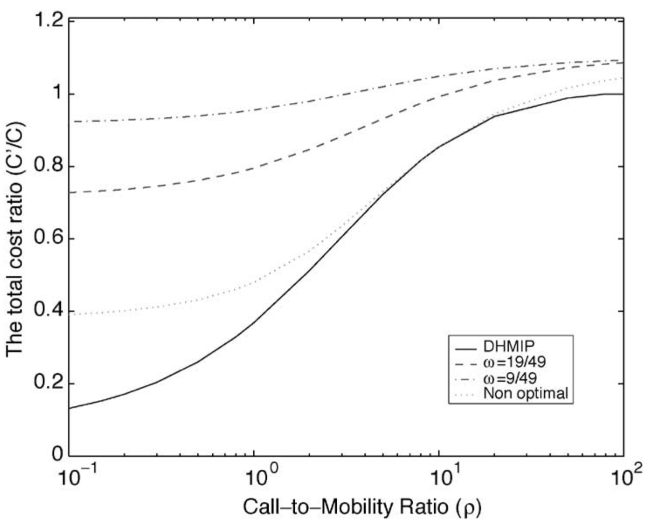



Fig. 10. Comparison with the IETF hierarchical scheme.

of subnets under one GFA is large, the HMIP scheme can reduce the system signaling traffic more when the CMR is small. However, the total cost increases rapidly with the increase of CMR. So, for the IETF HMIP scheme, it is hard to find an optimal number of subnets under one GFA for all the users in one network. Another drawback of the HMIP scheme, as mentioned in [11], is that the centralized system architecture makes the system performance sensitive to the failure of GFAs. Our DHMIP scheme avoids this problem successfully.

## VII. Some Improvements for the DHMIP Scheme

The performance of the DHMIP scheme can be improved further by the *IP paging* and *loop removal*.

### A. State Activation

Cellular IP supports IP paging [9]. In the IP paging protocol, the HA only keeps approximate location information of MH which is in idle state. The MH will turn to active state when packets are received and update its current location to the HA. We can enhance the performance of the DHMIP scheme by setting up a long threshold for MH in idle state in a similar way. In this paper, we name the improved DHMIP scheme *enhanced DHMIP scheme*. We define the *idle MH* as one that has not received data packet for a system predefined period. In the enhanced DHMIP scheme, the MH can keep a fixed relatively large threshold $K$ so that the MH can avoid updating the home network often. An idle MH that receives the first packet turns from idle to active state and updates its current location to the HA immediately and keeps doing it when it enters new subnets if it keeps in active state. In the enhanced DHMIP scheme, the data packets can avoid the extra transmission delay in the DHMIP scheme and the total signaling traffic is reduced at the same time. If the mobile user has not received packet for some predefined period, the MH returns to the idle state again. In this enhanced version, an MH can maintain a fixed threshold $K$ without computing and changing it from time to time. This can also reduce the mobile terminal energy consumption. For different mobile users with different call to mobility pattern, they

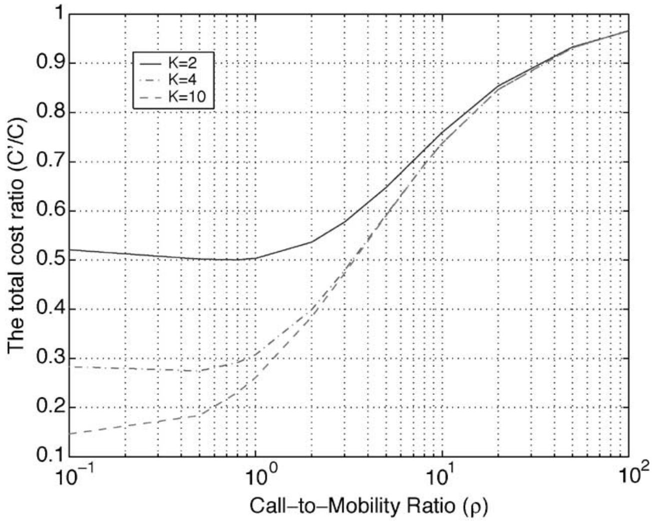



Fig. 11. Total cost for the enhanced DHMIP scheme.

can adopt different predefined thresholds. In our analysis, we assume that the MH turns to active state after receiving the first data packet and can update to its home network by attaching the location information to the acknowledge messages. Then, we can obtain the location update cost function $(EM')$, the packet delivery cost function $(EF')$ and the total cost function $(EC')$, respectively, as follows:

$$\begin{aligned} EM'(K,\rho) &= \sum_{i=0}^{\infty}\left[\left\lfloor\frac{i}{K}\right\rfloor U+\left(i-\left\lfloor\frac{i}{K}\right\rfloor\right)S\right]\alpha(i), \\ &= \frac{S}{\rho}+\frac{(1-g)g^{K-1}(U-S)}{\rho(1-g^K)} \end{aligned} \tag{24}$$

$$\begin{aligned} EF'(K,\rho) &= F+\sum_{i=0}^{\infty}\left(i-K\left\lfloor\frac{i}{K}\right\rfloor\right)T\alpha(i) \\ &= F+\frac{\left[1-Kg^{K-1}+(K-1)g^K\right]T}{\rho(1-g^K)} \end{aligned} \tag{25}$$

$$\begin{aligned} EC'(K,\rho) &= EM'(K,\rho)+EF'(K,\rho) \\ &= F+\frac{S}{\rho} \\ &\quad +\frac{(U-S)(1-g)g^{K-1}+\left[1-Kg^{K-1}+(K-1)g^K\right]T}{\rho(1-g^K)} \\ &= F+\frac{S+T}{\rho}+\frac{(U-S-KT)(1-g)g^{K-1}}{\rho(1-g^K)}. \end{aligned} \tag{26}$$

We further assume that the MH's FA residence time in subnets is exponentially distributed, then $g=1/(1+\rho)$. We can rewrite (26) as

$$EC'(K,\rho)=F+\frac{T+S}{\rho}+\frac{U-S-KT}{(1+\rho)^K-1}. \tag{27}$$

Fig. 11 shows the performance of the enhanced DHMIP scheme. We can observe from the figure that there is no optimal threshold for an MH. The larger the threshold, the better the performance. However, a large threshold may generate long packet delivery delay for the first data packet so that the network operator should choose the value by considering the signaling saving and the QoS comprehensively. In Fig. 11, we also see that the $K$ has little effect on the performance when the CMR is large. The reason is that the MH is always in active state and keeps updating its newly obtained CoA to the HA in this situation and the effective threshold reduces to one no matter what the predefined value is. In the enhanced DHMIP scheme, the total system signaling cost will never exceed that of the MIP scheme under all the conditions.

### B. Loop Removal

The performance of the DHMIP scheme can also be improved by removing the *loop* formed during the mobile user's movement. In reality, many mobile users roam in a limited region, for example, in an office building. If the mobile users revisit some subnets they have visited before, loop may form in the hierarchical architecture in the DHMIP scheme according to the protocol described before. With the *loop removal*, the MHs can update their new CoAs less often and the total signaling cost can be mitigated further [16]. When an MH enters a new subnet and tries to set up a new level of hierarchy, the new FA checks its hierarchy list first. If the FA is already in the MH hierarchy, the FA can delete the subsequent FA addresses for that user without updating the old FA so that the loop is removed. In this section, we analyze the *loop removal* effect on the DHMIP performance.

If we consider the MH revisiting scenario, we can define $\theta(i)$ as the effective number of subnets crossed between two consecutive packet arrivals. Although the *loop removal* can reduce the number of HA updates, the MH has to update its location to the current FA even the MH has visited that FA before. We define $\delta=2m_{mf\_u}+2p_{mf\_u}$ as the signaling cost of an MH notifying its arrival to the current FA, we can obtain the cost function with *loop removal* as

$$\begin{aligned} NC'(\kappa,K,\rho) &= \sum_{\kappa=0}^{K-1}\beta_\kappa \\ &\quad\times\left\{\sum_{i=0}^{\infty}\left[\left\lfloor\frac{\theta(i)+\kappa}{K}\right\rfloor(U-\delta)\right.\right. \\ &\qquad +\left(\theta(i)-\left\lfloor\frac{\theta(i)+\kappa}{K}\right\rfloor\right)(S-\delta)+i\delta \\ &\qquad \left.\left.+\left(\theta(i)+\kappa-\left\lfloor\frac{\theta(i)+\kappa}{K}\right\rfloor K\right)T\right]\alpha(i)+F\right\} \\ &= F+\frac{\delta}{\rho}+T\sum_{\kappa=0}^{K-1}\beta_\kappa\kappa \\ &\quad +(U-S-KT)\sum_{\kappa=0}^{K-1}\sum_{i=0}^{\infty}\beta_\kappa\left\lfloor\frac{\theta(i)+\kappa}{K}\right\rfloor\alpha(i) \\ &\quad +(S+T-\delta)\sum_{i=0}^{\infty}\theta(i)\alpha(i). \end{aligned} \tag{28}$$

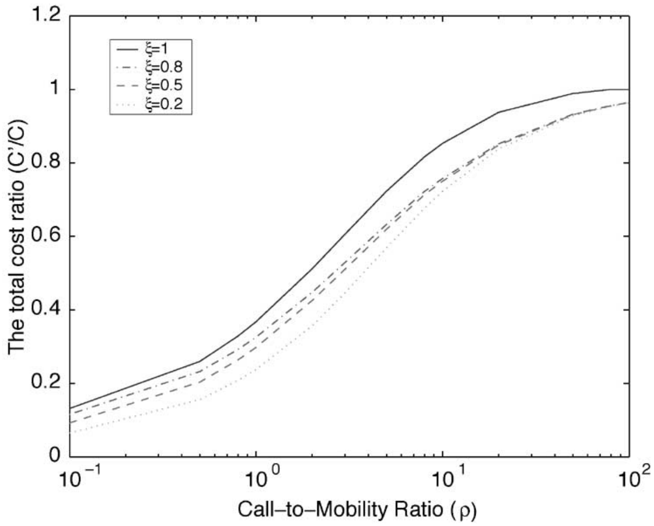



Fig. 12. Total cost with *loop removal*.

For simplicity and demonstration purpose, we assume that $\kappa$ is uniformly distributed and $\theta(i) = \xi i$, where $\xi$ is the percentage of the *effective* number of subnets a user visited to the total number of subnets the user visited. Then, we have

$$NC'(K,\rho) = F + \frac{(K-1)T}{2} + \frac{(S+T)\xi + (1-\xi)\delta}{\rho} + \frac{(U-S-KT)}{K} \sum_{\kappa=0}^{K-1} \sum_{i=0}^{\infty} \left\lfloor \frac{\xi i + \kappa}{K} \right\rfloor \alpha(i). \quad (29)$$

Fig. 12 shows the *loop removal* effect when $\xi$ is 1, 0.8, 0.5, and 0.2, respectively. When $\xi = 1$, it is the worst case that the mobile user never revisits any subnet he/her visited before. From Fig. 12, we can see that the total signaling cost for the DHMIP scheme can be reduced by the *loop removal* mechanism.

## VIII. Conclusion

This paper has provided a new location management scheme for MIP network—the DHMIP strategy. Instead of updating the home networks far away, in the DHMIP scheme, the MHs inform their new CoAs to the previous FAs (or one-hop router). When data packets arrive at the MHs, the packets can be delivered through the FA hierarchy. The transmission distance between the FAs is usually shorter than that between an MH and the HA, so that total signaling cost can be reduced. In this paper, we also proposed an iterative algorithm to compute the optimal threshold values for specific users with different call-to-mobility patterns. Our analysis shows that the DHMIP scheme outperforms the traditional MIP scheme and IETF HMIP protocol under various conditions. We also demonstrate that the DHMIP can further be improved by considering *state activation* and *loop removal*.

## References


[1] I. F. Akyildiz, J. McNair, J. S. M. Ho, H. Uzunalioglu, and W. Wang, "Mobility management in next-generation wireless systems," *Proc. IEEE*, vol. 87, pp. 1347–1384, Aug. 1999.

[2] F. M. Chiussi, D. A. Khotimsky, and S. Krishnan, "Mobility management in third-generation all-IP networks," *IEEE Commun. Mag.*, pp. 124–135, Sept. 2002.

[3] C. E. Perkins, "Mobile IP," *IEEE Commun. Mag.*, vol. 35, pp. 84–99, May 1997.

[4] C. Perkins, IIP mobility support for IPv4, in Internet RFC 3344, work in progress, Aug. 2002.

[5] R. Caceres and V. N. Padmanabhan, "Fast and scalable handoffs for wireless internetworks," in *Proc. ACM MOBICOM'96*, 1996, pp. 56–66.

[6] S. Das, A. Misra, P. Agrawal, and S. K. Das, "TeleMIP: telecommunications-enhanced mobile IP architecture for fast intradomain mobility," *IEEE Pers. Commun.*, vol. 7, pp. 50–58, Aug. 2000.

[7] A. T. Campbell, J. Gomez, S. Kim, C.-Y. Wan, Z. R. Turanyi, and A. G. Valko, "Comparison of IP micromobility protocols," *IEEE Wireless Commun.*, vol. 9, pp. 72–82, Feb. 2002.

[8] R. Ramjee, K. Varadhan, L. Salgarelli, S. R. Thuel, S.-Y. Wang, and T. La Porta, "HAWAII: a domain-based approach for supporting mobility in wide-area wireless networks," *IEEE/ACM Trans. Networking*, vol. 10, pp. 396–410, June 2002.

[9] A. T. Campbell, J. Gomez, and A. G. Valko, "An overview of cellular IP," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC'99)*, New Orleans, LA, Sept. 1999, pp. 606–610.

[10] E. Gustafsson, A. Jonsson, and C. Perkins, Mobile IPv4 Regional Registration, draft-ietf-mobileip-reg-tunnel-06.txt, work in progress, Mar. 2002.

[11] J. Xie and I. F. Akyildiz, "A distributed dynamic regional location management scheme for mobile IP," in *Proc. IEEE INFOCOM*, 2002, pp. 1069–1078.

[12] C. Perkins and K.-Y. Wang, "Optimized smooth handoffs in Mobile IP," in *Proc. IEEE Symp. Computers Communications*, Red Sea, Egypt, July 1999, pp. 340–346.

[13] C. Perkins and D. B. Johnson, Route Optimization in Mobile IP, draft-ietf-mobileip-optim-11.txt, work in progress, Sept. 2001.

[14] W. Ma and Y. Fang, "Location management employing two-level forwarding pointers in PCS networks," in *Proc. IEEE GLOBECOM'02*, Taipei, Taiwan, Nov. 2002, pp. 1653–1657.

[15] Y. Fang, I. Chlamtac, and Y. B. Lin, "Portable movement modeling for PCS networks," *IEEE Trans. Veh. Technol.*, vol. 87, pp. 1347–1384, Aug. 1999.

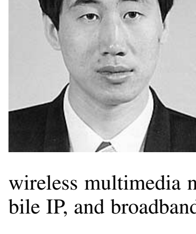
[16] R. Jain and Y. B. Lin, "An auxiliary user location strategy employing forwarding pointers to reduce network impacts of PCS," *Wireless Networks*, vol. 1, pp. 197–210, 1995.

[17] H. Xie, S. Tabbane, and D. J. Goodman, "Dynamic location area management and performance analysis," in *Proc. 43rd IEEE VTC*, 1993, pp. 536–539.

[18] J. S. M. Ho and I. F. Akyildiz, "Mobile user location update and paging under delay constraints," *ACM-Baltzer J. Wireless Networks (WINET)*, vol. 1, no. 4, pp. 413–425, Dec. 1995.

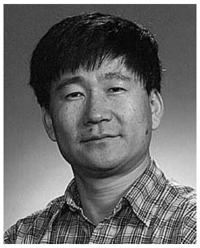
[19] Y. Bejerano and I. Cidon, "An anchor chain scheme for IP mobility management," *Wireless Networks*, vol. 9, no. 5, pp. 409–420, 2003.

[20] C.-H. Chu and C.-M. Weng, "Pointer forwarding MIPv6 mobility management," in *Proc. GLOBECOM*, Taipei, Taiwan, 2002, pp. 2133–2137.

**Wenchao Ma** (S'99–M'04) received the B.S. and M.S. degrees from Beijing University of Posts and Telecommunications, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 2003.

From 1998 to 1999, he was an Engineer with China Telecom. From 1999 to 2000, he was with the Electrical and Computer Engineering Department, New Jersey Institute of Technology, Newark. He is currently with Microsoft Research Asia, Beijing, as an Associate Researcher. His research interests include wireless multimedia networks, mobility management, mobile computing, mobile IP, and broadband wireless access.



**Yuguang (Michael) Fang** (S'92–M'94–SM'99) received the B.S. and M.S. degrees in mathematics from Qufu Normal University, Qufu, Shandong, China, in 1984 and 1987, respectively, the Ph.D. degree from the Department of Systems, Control and Industrial Engineering, Case Western Reserve University, Cleveland, OH, in January 1994, and the Ph.D. degree from the Department of Electrical and Computer Engineering, Boston University, Boston, MA, in May 1997.

From 1987 to 1988, he held research and teaching positions in both the Department of Mathematics and the Institute of Automation at Qufu Normal University. He held a postdoctoral position in the Department of Electrical and Computer Engineering, Boston University, from June 1994 to August 1995. From June 1997 to July 1998, he was a Visiting Assistant Professor in the Department of Electrical Engineering, University of Texas, Dallas. From July 1998 to May 2000, he was an Assistant Professor in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark. From May 2000 to July 2003, he was an Assistant Professor in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where he has been an Associate Professor since August 2003. His research interests span many areas including wireless networks, mobile computing, mobile communications, automatic control, and neural networks. He has published over 90 papers in refereed professional journals and conferences. He has actively engaged in many professional activities.

Dr. Fang is an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, an Editor for *ACM Wireless Networks*, an Area Editor for *ACM Mobile Computing and Communications Review*, an Associate Editor for *Wiley International Journal on Wireless Communications and Mobile Computing*, and an Editor for the the IEEE WIRELESS COMMUNICATIONS. He was an Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS: Wireless Communications Series, and the Feature Editor for *Scanning the Literature in IEEE Wireless Communications* (formerly *IEEE Personal Communications*). He has been actively involved with many professional conferences such as ACM MobiCom'02, ACM MobiCom'01, IEEE INFOCOM'04, INFOCOM'03, INFOCOM'00, INFOCOM'98, IEEE GLOBECOM'04 (Co-Chair for Global Internet and Next-Generation Networks Symposium), GLOBECOM'03, GLOBECOM'02, IEEE ICC'04, IEEE WCNC'04, WCNC'02, WCNC'00 (Technical Program Vice-Chair), WCNC'99, and International Conference on Computer Communications and Networking (IC3N'98) (Technical Program Vice Chair). He received the National Science Foundation Faculty Early Career Development Award in 2001 and the Office of Naval Research Young Investigator Award in 2002. He is listed in Marquis *Who's Who in Science and Engineering*, *Who's Who in America*, and *Who's Who in World*. He is a member of the Association for Computing Machinery (ACM).