# P-MIP: Paging in Mobile IP


Xiaowei Zhang
Comet Group
Columbia University
New York, NY 10027
xzhang@comet.columbia.edu

Javier Gomez Castellanos
Comet Group
Columbia University
New York, NY 10027
javierg@comet.columbia.edu

Andrew T. Campbell
Comet Group
Columbia University
New York, NY 10027
campbell@comet.columbia.edu



## ABSTRACT

*As the number of Mobile IP users grows, so will the signaling overhead associated with Internet mobility management in the core IP network. This presents a significant challenge to Mobile IP as the number of mobile devices scale-up. In cellular networks, registration and paging techniques are used to minimize the signaling overhead and optimize the mobility management performance. Currently, Mobile IP supports registration but not paging. In this paper, we argue that Mobile IP should be extended to support paging to improve the scalability of the protocol to handle large populations of mobile devices. To address this, we introduce P-MIP, a set of simple paging extensions for Mobile IP, and discuss the construction of paging areas, movement detection, registration, paging and data handling. We present analysis and simulation results for Mobile IP with and without paging extensions, and show that P-MIP can scale well supporting large numbers of mobile devices with reduced signaling under a wide variety of system conditions. The ns-2 P-MIP source code used to evaluate the protocol in this paper is available on the web (comet.columbia.edu/pmip).*


## 1. INTRODUCTION

Mobile IP (MIP) [1] represents a global mobility solution that provides host mobility support for a wide variety of radio technologies, devices, and applications. Mobile IP supports a simple mobility mechanism. A mobile node's location is tracked by its home agent which binds the care-of address (CoA) or co-located care-of address used by the mobile node in a visited network to the mobile node's home address. When a mobile node moves to a new network, it "registers" its new care-of address with its home agent. This basic mechanism presents a number of challenges to the widespread deployment of Mobile IP with respect to handoff performance and scalability issues. In this paper, we address Mobile IP scalability in support of very large numbers of mobile nodes.

Currently, mobility management is supported by mobile node registration in Mobile IP. Considering the fact that Mobile IP users will be idle most of time, it is not necessary to keep tracking their exact location all the time. We argue in this paper that as wireless access to Internet becomes the norm, Mobile IP will need to provide a more scalable and efficient location tracking scheme that distinguishes between mobile nodes that are actively communicating or idle. In this case, it would be sufficient for Mobile IP only to know the approximate location of idle users to resolve this scalability issue. The exact location of idle mobile nodes can be found by using *paging*.

Paging is a procedure that allows a wireless system to search for an idle mobile host when there is a message destined to it, such that the mobile user does not need to register its precise location to the system whenever it moves. Paging has a number of benefits. First, it reduces signaling overhead associated with registration and location system database updates. The power consumption of mobile nodes is also reduced because idle mobile nodes no longer have to register their exact location with the system. Paging, however, introduces uncertainty about the location of mobile hosts, and as such, can result in additional delays associated with message delivery. Paging also introduces additional signaling to the system that consumes system resources. Thus, it is important to balance the paging and registration processes in order to optimize system performance.

The contribution of this paper is as follows. We address Mobile IP scalability and propose P-MIP [2], paging extensions for Mobile IP. We show that the addition of a set of simple paging extensions can provide significant savings in signaling overhead, helping Mobile IP to scale to support large numbers of mobile users. This paper is structured as follows. Section 2 discusses related work and Section 3 provides an overview of P-MIP and its operations. Section 4 presents the detailed design of the protocol, which assumes the use of foreign agents. We discuss P-MIP's support for paging area construction, movement detection, registration, paging and data handling. In Section 5, we evaluate the signaling cost of Mobile IP with and without the P-MIP paging extensions. We study how paging area size, mobile speed, mobile density and data session characteristics can impact the signaling cost under a variety of system conditions. We use a combination of analysis and simulation to evaluate system performance and present our results. Finally, in Section

6 we make some concluding remarks.

## 2. RELATED WORK

Paging is used widely in cellular systems [4] [5] to locate idle mobile nodes prior to establishing incoming calls. A paging area is constructed by a group of base stations and, typically, these base stations are under the control of the same mobile switching center (MSC). When an incoming call for a mobile user is received, the mobile switching center sends a paging message to all the base stations in the same paging area. Each base station broadcasts the paging message in its own cell. The system determines the mobile station's accurate location after receiving a paging response message from the paged mobile user. The precise location information is then used to establish the call.

Cellular networks are connection-oriented and paging is only necessary in support of incoming calls. Mobile IP, however, is based on a connectionless service paradigm where datagrams are treated independently from one another. The interpretation of paging in IP networks is therefore an open issue. Clearly, paging would represent a considerable overhead if mobile hosts were paged on each incoming packet. This issue is resolved in P-MIP by using mobile host state and the concept of a data session. Mobile node state consists of a per-node tuple (operational mode, active state timer) where mobiles can be in an *active* or *idle* operational mode. A mobile node is in active mode if it has recently sent or received IP data. A mobile node is considered to be active for an active timer period starting from the instance the node sent or received data. Each time the mobile node sends or receives data, the timer is reset in both the mobile node and the serving foreign agent. When the active timer expires, the mobile node enters idle state. Whenever an idle node sends or receives data, its operational mode is changed to active and its active timer started. The active state timer value is implementation dependent. If packets in a packet train are close enough, then a mobile node state will remain active during the packet train transmission. In P-MIP, we call this a *data session.* We can broadly consider the "holding time" of data sessions as analogous to call holding times found in connection-oriented networks. However, in some cases (e.g., Web interactions), these are very short-lived sessions (e.g., microflows). As a general design rule, P-MIP only needs to page a mobile node on the first packet of a data session.

Recently, a number of micro-mobility protocols (e.g., Cellular IP [6], Hawaii [8] and Regional Registration [9]) have been primarily developed to support better handoff performance in environments where mobile hosts handoff frequently. Micro-mobility protocols reduce registration signaling and improve performance by minimizing delay and packet loss during handoff. However, these fast handoff approaches do not improve the power consumed by mobile nodes or the bandwidth consumed over the air interface. In contrast, paging does not reduce the registration delay. Rather, paging reduces the number of registrations that mobile nodes make. Paging can reduce the power consumption of mobile nodes, and possibly the bandwidth used over the air interface. In this paper, we focus on paging and not fast handoff control. By extending Mobile IP to support paging, P-MIP can reduce the signaling load in the core IP network, power consumption of mobile nodes, and potentially radio resources.

Two micro-mobility protocols, namely Cellular IP [6] and Hawaii [8], integrate support for fast handoff and paging. These protocols, which operate in wireless access networks and interwork with Mobile IP via mobile gateways, represent independent protocols that are rather different in comparison to the base Mobile IP protocol [1]. We take a different approach and propose P-MIP, a simple set of paging extensions to the base Mobile IPv4 protocol.

In [10], paging is added as an extension to both Mobile IP [1] and regional registrations [9]. Here mobile hosts need to register using an Idle Mode Request extension to enter idle mode. In P-MIP, the modes of mobile hosts are controlled by timers which can reduce registeration signaling even further. In addition, P-MIP does not depend on the regional registrations protocol [9].

In [11], Mobile IP is extended with an adaptive paging scheme. In this approach, mobile nodes support optimum paging areas, where a paging area is adaptive on a per-mobile basis. Mobile nodes continuously compute optimal paging areas. This operation can, however, adversely impact the power consumption of mobile nodes. In addition, some of the input parameters used by the optimization process are difficult for a mobile node to determine. The adaptive paging scheme also calls for protocol modifications that are incompatible with Mobile IP. P-MIP defines a simple set of paging extensions to the base Mobile IP protocol and is backward compatible with Mobile IP. We believe P-MIP represents a more synergistic approach to the base protocol for the delivery of value-added Mobile IP paging services.

## 3. P-MIP OVERVIEW

In what follows, we present an overview of P-MIP. P-MIP extends the Mobile IP [1] agent advertisement message to include a "P bit". Foreign agents supporting paging set the P bit in agent advertisements. As a default, however, the P-bit is not set. P-MIP assumes the use of foreign agents and that mobile nodes register through foreign agents[1]. When a mobile node first receives an agent advertisement, it checks if paging is supported by looking at the "P bit". Note that there is also a P bit in the registration request message, which is used to indicate if mobile nodes support paging or not. In this paper, we assume both foreign agents and mobile nodes support paging.

An active mobile node operates in exactly the same manner as in Mobile IP [1]. When a mobile node changes its point of attachment, it registers. When an idle mobile node moves to a new paging area, it registers. Idle mobile nodes do not register when moving within a paging area.

When packets are destined to mobile nodes, home agents forward them to the foreign agents registered by the mo-

[1]Future work will consider paging in environments that do not support foreign agents, (e.g., when collocated care-of addresses are used).

bile nodes, known as registered foreign agents. A registered foreign agent first determines if it has any information on record[2] for the mobile node. If a record exists, then the registered foreign agent checks if the mobile node supports paging or not. If the mobile node supports paging, then the registered foreign agent checks the mobile node's operational state. If the mobile node is in active mode, then the registered foreign agent decapsulates packets and forwards them to the mobile node, as in the case of Mobile IP [1]. If the mobile node is in idle mode, then the registered foreign agent sends a *paging request* message to all other foreign agents that reside in the same paging area, as well as broadcasts the paging request to its own network.

We use the following terminology to clarify the interaction between system components that implement P-MIP. The foreign agent serving the network that a mobile host is currently visiting is called the *current foreign agent*; the foreign agent a mobile host registers with is called the *registered foreign agent* which may or may not be on the same network that the mobile node is currently located; the foreign agent that originates the paging request is called the *paging foreign agent*; and a mobile node that is paged is called the *paged mobile node*.

When a mobile node receives a paging request message, it registers with its home agent through the foreign agent on the visiting network (i.e., the current foreign agent). After receiving a registration reply message, the mobile node sends a paging reply message back to the foreign agent it had previously registered with through its current foreign agent to inform the previously registered foreign agent of its current location. Note that a mobile node's current foreign agent and its previously registered foreign agent are the same foreign agent if it has not moved out of the network since the last registration. When the previously registered foreign agent receives a paging reply message, it forwards any buffered packets toward the mobile node through its current foreign agent.

## 4. P-MIP DESIGN

The design of P-MIP encompasses paging area construction, movement detection, registration, paging and data handling.

### 4.1 Paging Area Construction

In P-MIP, a paging area is identified as having a unique paging area ID (PAI) and consists of two or more networks, each of which is identified by a network prefix as part of its Internet address. Paging areas can be configured based on a number of criteria (e.g., node mobility, traffic patterns, mobile density, etc.) so that most mobile nodes move within the same paging area. Paging areas can be configured manually by administrators by setting parameters at each foreign agent or more automatically by having foreign agents interact with paging servers. The benefit of paging servers is that administrators only need to configure the servers. In this case, each foreign agent acquires its paging area information directly from a paging server.

We consider two types of paging area construction in P-MIP that encompass *non-overlapping* and *overlapping paging areas*[3]. In the case of non-overlapping paging areas, a network can only be associated with one paging area, as shown for paging areas PA1, PA2, and PA3 in Figure 1. Note that the figure does not show any specific connections within or between paging areas. A number of possible connectivity scenarios can be considered. For example, all base stations in a paging area are directly connected to the Internet, or all access networks in a paging area are connected to the Internet through one or more IP gateways.

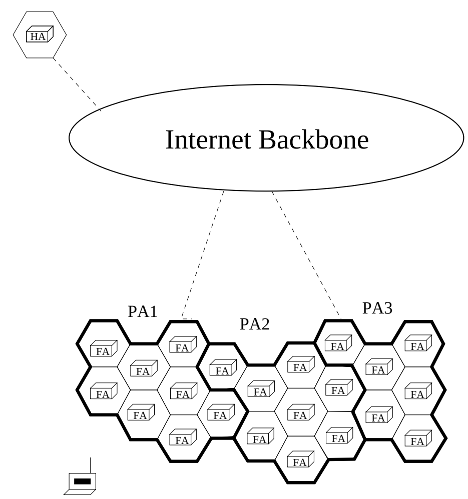


**Figure 1: Non-overlapping Paging Areas**

In the non-overlapping scheme, each paging area has a unique paging area ID. Paging area IDs can be reused in the same sense as spatial reuse with adjacent paging areas having different IDs. Paging area ID assignment could follow the same lines as IP addresses, with authorized organizations managing allocation and assignment.

A paging table is maintained by each foreign agent and can be manually set or configured by a paging server or an administrator. In P-MIP, the agent advertisement message is extended to carry the paging area ID. The agent advertisement message must be broadcast periodically so that the paging area ID can be periodically distributed. The format of the paging tables for the non-overlapping scheme is shown in the following table.

[2]The term "record" used in this paper refers to Mobile IP mobility management related information maintained by home and foreign agents and mobile nodes. A record comprises transient state information including the mobile node's home address, home agent address, care-of address, and various timer related information associated with control messaging (e.g., registration lifetime, etc.).

[3]Our approach to overlapping paging area construction is similar to the work on adaptive paging areas [11]. In P-MIP, however, each paging area is based on a foreign agent rather than a mobile node. In addition, P-MIP paging areas are preconfigured and fixed which simplifies the protocol, rather than being adaptive [11].

| PAI | XXXX XXXX |
|---|---|
| FA1 | FA1's IP address |
| FA2 | FA2's IP address |
| FA3 | FA3's IP address |
| . | . |
| . | . |
| . | . |

In the overlapping paging areas scheme, each foreign agent is located at the center[4] of its paging area. As shown in Figure 2, foreign agent FA1 is located at the center of paging area PA1, FA2 is located at the center of PA2, and so on. In this paper, we concentrate on the non-overlapping scheme.

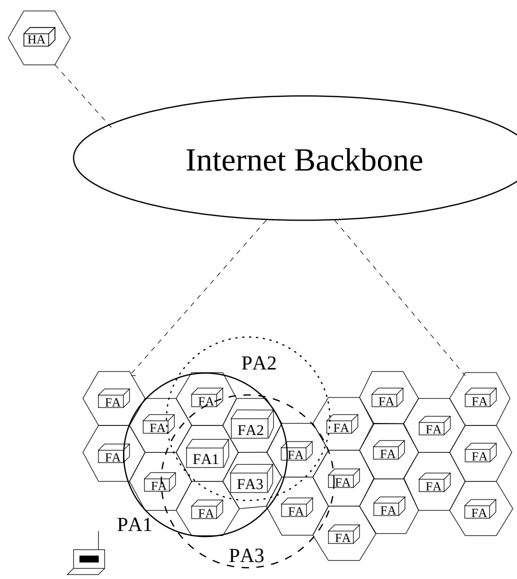


**Figure 2: Overlapping Paging Areas**

## 4.2 Movement Detection

The algorithms used to detect whether a mobile node has moved to a different network are the same for P-MIP as Mobile IP. Two algorithms are used for movement detection in Mobile IP. In the first approach, mobile nodes detect movement using agent advertisement messages and the advertisement lifetime. When a mobile node receives an agent advertisement, it extracts the advertisement lifetime included in the message. If a mobile node does not receive another advertisement before the advertisement lifetime expires, then the mobile node assumes that it has moved to a new network. In the second approach, a mobile node compares the network prefix obtained from newly received agent advertisement messages to the current care-of address. If they are different, the mobile node assumes it has moved to another network.

If a mobile node is in an active mode when it detects it has moved, then handoff is initiated which invokes the registration process. In this case, the mobile node's location will be updated at both the foreign and home agents. In contrast, if a mobile node is in an idle state and it detects it has moved to a new network within the same paging area, then registration is unnecessary. However, if an idle mobile node does not have a valid foreign agent on record, then it should locate one for further movement detection and registration purposes.

The algorithm used to detect movement between paging areas is as follows: a mobile node records the advertisement lifetime from the agent advertisement message. If a mobile node fails to receive another advertisement from the same agent within the specified lifetime, then the mobile node assumes that it has lost contact with its registered foreign agent. If the mobile node has previously received an agent advertisement from another foreign agent for which the lifetime field has not yet expired, then the mobile node should check if the paging area ID from the new agent is the same as its current paging ID. If the paging IDs are the same, then the mobile node will do nothing except track the foreign agents identified by agent advertisement messages. If the paging IDs are different, then idle mobile nodes should immediately register with the new foreign agent. If the mobile node is not in contact with any foreign agent, then it sends an agent solicitation message. Upon reception of an agent advertisement message, the mobile node checks if the new paging area ID is the same as the old ID. The mobile node uses this test to determine whether it has moved to a new paging area or not.

## 4.3 Registration

Registration in P-MIP follows the same procedure as Mobile IP. However, after receiving a registration reply message in P-MIP, both the foreign agent and the mobile node update the mobile node state by setting the operational mode to active and starting the active timer. In the case overlapping paging areas, the registration reply message includes the paging area ID. The list of foreign agents in the paging area is inserted by the foreign agent in the registration reply message returned to the mobile node. The mobile node records the paging area ID.

An important performance difference between Mobile IP and P-MIP relates to the number of registration signaling messages generated. In Mobile IP, a mobile node registers when a mobile node changes its point of attachment, when a mobile node's registration lifetime is about to expire (we call this "registration refresh"), and when a mobile node detects that its current foreign agent has rebooted. In P-MIP, a mobile node registers when a mobile node detects it has moved to a new paging area, or when a mobile node's registration lifetime is about to expire, or when a mobile nodes detects that its current foreign agent has rebooted, or when a mobile node is paged and is not on the same network as the paging foreign agent, or when an idle mobile node is about to transmit data.

Because idle mobile nodes do not register when moving between networks in the same paging area, P-MIP reduces

[4]By the term "center" we mean the optimum position where a foreign agent is positioned in its paging area such that a mobile node in this network minimizes its chance of moving out of the current paging area. In this manner, the signaling generated by boundary crossing can be reduced to a minimum.

signaling in comparison to the base Mobile IP protocol. We discuss this issue when we present our performance evaluation in Section 5.

### 4.4 Paging

When a foreign agent has to locate an idle mobile node, it broadcasts a paging request message to its own network and to all other foreign agents in its paging area. We use unicast[5] as a basis for broadcasting paging messages to all other foreign agents in the paging area. Foreign agents broadcast the paging request within their own wireless networks. P-MIP supports the optimization that a foreign agent may have a number of paging requests to make at any one moment. In this case, P-MIP aggregates all paging requests into a single paging message. This helps to minimize the paging overhead as the number of paged mobile nodes grows. The paging request message lists the home addresses of all paged mobile nodes. In this paper, we focus on the scenario where a foreign agent pages a single mobile node.

When a foreign agent receives a paging request from the paging foreign agent, it first conducts an authentication check if the foreign agents share a mobility security association[6]. If authentication is successful, then the foreign agent broadcasts the paging message on its own network. When a mobile node finds its home address in the paging request and the message is directly from its registered foreign agent, then the registered foreign agent is the same one as the mobile host's current foreign agent. Thus, the mobile node deduces that it is still located in the same network as its registered foreign agent. In this case, the mobile node sends a paging reply back to the paging foreign agent without registering, and sets its operational mode to active and restarts its active timer. Otherwise, the mobile node starts the registration procedure to make sure that the binding between its home address and new care-of address is valid. Following this, the mobile node responds to the paging request by sending a paging reply. The mobile node will insert its care-of address in the paging reply. The transmission time for each registration and paging reply to a foreign agent is randomized to avoid potential collisions between signaling messages (e.g., registrations and paging reply messages) from multiple paged mobile nodes in the same network.

When the foreign agent receives the registration reply from a home agent, it sets the mobile node's operational mode to active and then relays the registration reply to the mobile node. If the mobile node does not receive the registration reply, its mode remains idle in its own record. Loss of the registration reply message can result in state inconsistency between the foreign agent and mobile node records. However, this state inconsistency is short lived. When the foreign agent receives the registration reply message, it starts to forward data to the mobile node. If a mobile node receives data packets in idle mode, it checks if this foreign agent is the one to which it sent a registration request. If this is the case, then the mobile node sets its care-of address to the one associated with the foreign agent, changes mode to active, and then starts its active timer.

A foreign agent conducts authentication checks on paging reply messages received from paged mobile nodes if the foreign agent and the mobile nodes share a mobility security association. If the foreign agent did not originate paging, then it will forward the paging reply message to the paging foreign agent. The foreign agent takes this action after appending the foreign agent-to-foreign agent authentication extension to the paging reply message. If the paging foreign agent receives a paging reply message directly from a paged mobile node, then it sets the mobile node's operational state to active and starts the active timer. The paging foreign agent removes "records" for all paged mobile nodes that are not located in its network. The paging foreign agent can determine this from paging reply messages. Paging foreign agents take this action because paged mobile nodes located in other paging areas have already initiated registration through foreign agents in the new network.

The paging foreign agent will retransmit a paging request after a "paging period" if no response is received from an earlier paging request. If a mobile node's registration lifetime expires during paging, then the paging foreign agent stops paging. Each successive paging period is at least twice the previous duration. The foreign agent stops paging after sending multiple successive paging requests without receiving a paging reply. The specific number of paging request retransmissions are implementation dependent. However, the paging foreign agent maintains the paged mobile node's record until the lifetime expires. Maintaining the record in this manner counters the situation where mobile nodes are temporarily partitioned in the network.

In the case where a mobile node moves to a new network, loss of paging reply messages does not have any serious impact on the mobile node's operation. This is because a mobile node will register with a new foreign agent before sending out a paging reply message. Therefore, the record maintained by a home agent points to the current foreign agent. However, it may cause unnecessary data loss for the mobile node which is in the same network. In this case, the mobile node sends out a paging reply and sets the mode to active. If a paging reply message is lost, the mobile node's operational mode remains idle in the record maintained by the foreign agent. In this case, the foreign agent continues paging the mobile node. The paging foreign agent will stop paging after multiple attempts and discard any buffered data. The following action attempts to resolve this condition. After a mobile node (which is in the same network) sends a paging reply message, it assumes reception of data from the foreign agent. If this does not happen after a certain period, then the mobile node assumes that the paging reply has been lost and sends a paging reply again. If the mobile node still does not receive data after multiple paging

[5] IP Multicast would provide better performance than simple unicast distribution of paging request messages in paging areas. However, IP Multicast is rather complex to implement in wireless access networks. Therefore, we adopt unicast for simplicity at this stage of our work. Future work will consider more efficient support for the transport of signaling messages in paging area through the implementation of a lightweight multicast service tailored for such a task.

[6] The authentication method used in P-MIP is the same as in Mobile IP [1].

reply retransmission attempts, it resets its operational mode to idle. The paging foreign agent will forward any buffered data toward the mobile node once a paging reply has been received by the paging foreign agent.

## 4.5 Data Handling

When there are incoming packets destined to a mobile node, the mobile node's home agent forwards data to its care-of address associated with the mobile node's registered foreign agent. When the registered foreign agent receives data, it checks the mobile node's state. If the mobile node is in an active state, then it is likely that the mobile node is located in the registered foreign agent's network. In this case, the foreign agent decapsulates and forwards packets to the mobile node. As discussed earlier, the registered foreign agent sends a paging request message if the mobile node is in an idle state, and at the same time, buffers the data for the mobile node. When the paging foreign agent (which is also the registered foreign agent) receives a paging reply, it forwards the buffered data to the mobile node. If the mobile node is not in the registered foreign agent's network, then the registered foreign agent removes the mobile node's record.

A foreign agent restarts a mobile node's active timer each time a data packet is sent or received to or from the mobile node, respectively. The mobile host also restarts its active timer under these conditions. In the case that a mobile node does not respond to multiple page request messages, the registered foreign agent discards any buffered data.

A mobile node can simply send data when it is in active mode. If the mobile node is idle, it first determines if it is still located on the same network that it previously registered with by checking the agent advertisement or by sending an agent solicitation message. If the mobile node has moved to a new network, then it initiates a registration procedure to update its location record in its home agent and foreign agent before sending data. It is likely that a mobile node which is sending data (e.g., TCP segments) will also receive data (e.g., TCP acknowledgements) in return. Therefore, the home agent and foreign agent bindings need to be updated to facilitate the delivery of incoming data to the mobile node. If the mobile node resides on the registered foreign agent network, then it does not need to register. Rather, the mobile node only needs to set its operational mode to active and starts its active timer before transmitting data.

# 5. PERFORMANCE EVALUATION

We use a combination of analysis and simulation to evaluate the signaling cost of Mobile IP with and without the P-MIP paging extensions. We study how paging area size, mobile speed, mobile density, and session characteristics can impact the signaling cost under a variety of system conditions. The signaling cost is represented as the product of the weighted distance (see Section 5.2.1 for details) the signaling message travels and the signaling rate. The unit of cost is weighted hops*pkt/sec. For simplicity, the air interface is treated as a single hop in the wireline network. Because agent advertisement and solicitation signaling present the same costs in Mobile IP and P-MIP, we exclude them from the evaluation. Therefore, only registration and paging signaling are considered. Simulation is also used to investigate scenarios that cannot be considered by analysis, for example, using a range of mobile speeds other than a constant speed, which is considered in the analysis. P-MIP is implemented as extensions to the ns simulator [17] Mobile IP [16] code. Simulation is also used to validate the analysis. The ns-2 P-MIP source code used to evaluate the protocol is available on the web http://comet.columbia.edu/pmip.

## 5.1 Mobility Models

Paging aims to reduce the signaling overhead associated with mobility in wireless networks. The signaling analysis is therefore closely related to the movement patterns of mobile nodes. In legacy wireless networks where voice is the dominant service, the fluid flow model is widely used to analyze cell boundary crossing related issues, such as hand-off [12] [13]. For simplicity, we adopt this model to analyze the signaling cost in a Mobile IP environment. We also assume that paging areas and wireless cells are square-shaped where there are n cells in a paging area. The perimeter of a cell is $l$, so the perimeter of the paging area, denoted as $L$, is $L = l\sqrt{n}$. Mobile nodes move at an average velocity of $v$ in directions that are uniformly distributed over $[0, 2\pi]$ and are uniformly distributed with density $\rho$. The cell boundary crossing rate $r$ is

$$r_c = \frac{\rho v l}{\pi} \tag{1}$$

where:

$r_c$ is the cell crossing rate (mobiles/sec);

$\rho$ is the mobile density ($mobiles/m^2$);

$v$ is the moving velocity (m/sec); and

$l$ is the cell perimeter (m).

Mobile devices move across a boundary in two directions. For evaluation purposes, however, only one direction needs to be considered. The paging area boundary crossing rate $r_p$ is

$$r_p = \frac{\rho v L}{\pi} \tag{2}$$

One drawback of the fluid flow mobility model used in the analysis is that mobile nodes are assumed to move at fixed rates. The mobility model used in the simulation supports a wide variety of mobility behavior. The mobility model is based on the random waypoint algorithm [18]. A mobile node picks a random location in the simulation area as the destination and moves towards the location at a speed chosen uniformly between 0 and the maximum speed. When a mobile node reaches a destination point, it stops for the duration of pause time and then picks another destination and speed, and moves again. This cycle is repeated for the duration of the simulation. When the pause time is set to the simulation duration time, the mobile node remains stationary. As the pause time increases so does the movement of the mobile node. Continuous movement corresponds to a pause time of 0. Node mobility movement patterns are generated as in [17].

## 5.2 Signaling Cost Analysis

### 5.2.1 Mobile IP

The formula used to calculate the signaling cost for Mobile IP is

$$\begin{aligned} C &= d_{HA,FA}\left(R_{core}w_{core}+R_{local}w_{local}\right)\Big[r_c n \\ &\quad + \rho\left(\frac{l}{4}\right)^2 n r_r\Big] \\ &= d_{HA,FA}\left(R_{core}w_{core}+R_{local}w_{local}\right)\Big[\frac{\rho v l}{\pi} n \\ &\quad + \rho\left(\frac{l}{4}\right)^2 n r_r\Big] \end{aligned} \tag{3}$$

where:

$C$ is the signaling cost for Mobile IP (weighted hops*msg/sec);

$d_{HA,FA}$ is the average distance between home and foreign agent in terms of the number of hops;

$R_{core}$ is the ratio of the number of hops in the core network to the total number of hops between a home agent and a foreign agent;

$R_{local}$ is the ratio of the number of hops in local access networks to the total number of hops between a home agent and foreign agent;

$w_{core}$ is the weight of each hop in the IP core network;

$w_{local}$ is the weight of each hop in access networks;

$r_c$ , $\rho$, $v$, and $l$ are the same as those in equation (1);

$n$ is the number of cells considered in a paging area;

$r_r$ is the average mobile registration refresh rate, which is related to registration lifetime and registrations triggered for other reasons (e.g., by an agent advertisement.

The transport of signaling messages over hops in the IP core network has different impact on the signaling cost in comparison to transporting signaling messages in local wireless access networks. The reason for this is that the signaling load introduced in the core IP network will potentially affect other networks and users, while signaling load on the access networks only has local effect. This is represented as weight factors $w_{core}$ and $w_{local}$ in Equation (3). The hop weight relates to the hop length (i.e., distance), CPU packet processing time, queuing delay, etc. The hop ratios $R_{core}$ and $R_{local}$ indicate the percentage of hops in the core network and local access networks, respectively. The "local hops" include those hops located in both the foreign and home agent access networks.

In Equation (3), the first term in brackets represents the signaling cost due to mobility. The second term represents the signaling cost due to registration refresh.

### 5.2.2 P-MIP

The signaling cost associated with P-MIP includes registration and paging signaling, and is

$$\begin{aligned} C_p &= d_{HA,FA}\left(R_{core}w_{core}+R_{local}w_{local}\right)\Big[r_p \\ &\quad + \left(r_c n - r_p\right)\alpha + \rho\left(\frac{l}{4}\right)^2 n r_r \\ &\quad + \rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\left(\lambda_d+\lambda_a\right)\Big] \\ &\quad + (n-1)\, d_{FA,FA}\ w_{local}\ \rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\lambda_a \\ &= d_{HA,FA}\left(R_{core}w_{core}+R_{local}w_{local}\right)\Big[\frac{\rho v L}{\pi} \\ &\quad + \frac{\rho v l}{\pi}\left(n-\sqrt{n}\right)\alpha + \rho\left(\frac{l}{4}\right)^2 n r_r \\ &\quad + \rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\left(\lambda_d+\lambda_a\right)\Big] \\ &\quad + (n-1)\, d_{FA,FA}\ w_{local}\ \rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\lambda_a \end{aligned} \tag{4}$$

where:

$C_p$ is the signaling cost for P-MIP (weighted hops * msg/sec);

$d_{HA,FA}$ , $R_{core}$, $R_{local}$, $w_{core}$, $w_{local}$, $r_c$, $\rho$, $n$, $v$, and $l$ are the same as those in the fomula of Mobile IP.

$d_{FA,FA}$ is the average distance between foreign agents (hops);

$r_p$ is the paging area crossing rate (mobiles/sec);

$\alpha$ is the ratio of active mobile nodes to total number of mobile nodes;

$\lambda_a$ is the incoming data session rate for a mobile node, it is also the paging rate for a mobile node (1/sec);

$\lambda_d$ is the outgoing data session rate for an idle mobile node (1/sec);

$L$ is the paging area perimeter (m); and

$r_r$ is the average mobile registration signaling refreshing rate which is related to registration lifetime and registrations caused for other reasons.

The first long term, including every short terms inside the brackets, and second long term of Equation (4) represent the signaling cost associated with registration and paging. The short terms within the first long term represent the registration signaling cost caused by (i) crossing paging area boundaries; (ii) active mobile nodes crossing cell boundaries; (iii) registration refresh; and (iv), the registration signaling caused by incoming and outgoing data to and from idle mobile nodes, respectively.

For registration refresh, we may adjust the registration lifetime in P-MIP to make the refreshing rate in Mobile IP and P-MIP equivalent. Thus, in the signaling analysis, we ignore the effect of refresh signaling and focus on the signaling difference with and without paging. We normalize the signaling cost to the weight the distance between a home agent and foreign agent. Equation (3) and (4) reduce to

$$\begin{aligned} C_n &= \frac{C}{d_{HA,FA}\left(R_{core}w_{core} + R_{local}w_{local}\right)} \\ &= \frac{\rho v l}{\pi} n \end{aligned} \quad (5)$$

$$\begin{aligned} C_{pn} &= \frac{C_p}{d_{HA,FA}\left(R_{core}w_{core} + R_{local}w_{local}\right)} \\ &= \frac{\rho v l \sqrt{n}}{\pi} + \frac{\rho v l}{\pi}\left(n - \sqrt{n}\right)\alpha \\ &\quad + \rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\left(\lambda_d + \lambda_a\right) \\ &\quad + \frac{d_{FA,FA}\left(n-1\right)\rho\left(\frac{l}{4}\right)^2 n\left(1-\alpha\right)\lambda_a}{d_{HA,FA}\left(R_{core}R_w + R_{local}\right)} \end{aligned} \quad (6)$$

where the hop weight ratio $R_w$ (discussed in Section 5.3) is the ratio of $w_{core}$ to $w_{local}$.

## 5.3 Analysis Results

### 5.3.1 Paging Area Size.

In what follows, we show the relationship between the signaling cost and the number of cells in a paging area. Most parameters used in this evaluation are set to typical values found in the literature for analyzing cellular systems [14]:

| | |
|---|---|
| $R_w$ | 8 |
| $R_{core}$ | 0.5 |
| $R_{local}$ | 0.5 |
| $d_{HA,FA}$ | 16 (hops) |
| $d_{FA,FA}$ | $\sqrt{n}$ (hops) |
| $v$ | 65 (mph) i.e. 28.9 (m/s) |
| $\rho$ | 200 ($users/km^2$) i.e. 0.0002 ($users/m^2$) |
| $\alpha$ | 5% [6] |
| $\lambda_a$ | 3(1/hr) i.e. 0.0008 (1/sec) |
| $\lambda_d$ | 3(1/hr) i.e. 0.0008 (1/sec) |

The average number of hops between foreign agents, $d_{FA,FA}$, is dependent on the network topology of the paging area. For simplicity we assume the average number of hops to be $\sqrt{n}$. The number and ratio of hops, $d_{HA,FA}$, $R_{core}$, $R_{local}$, and $R_w$ were determined experimentally using the Internet traceroute tool to obtain an approximate picture of hop numbers and core/local hop distribution between Columbia University and other US universities. Typically, $w_{core}$ is greater than $w_{local}$, resulting in a hop weight ratio ($R_w$) greater than 1. However, to accurately determine the hop weight ratio is difficult. Here, we set $R_w$ to 8. The impact of different hop weight ratio $R_w$ on the signaling cost is discussed later in this section.

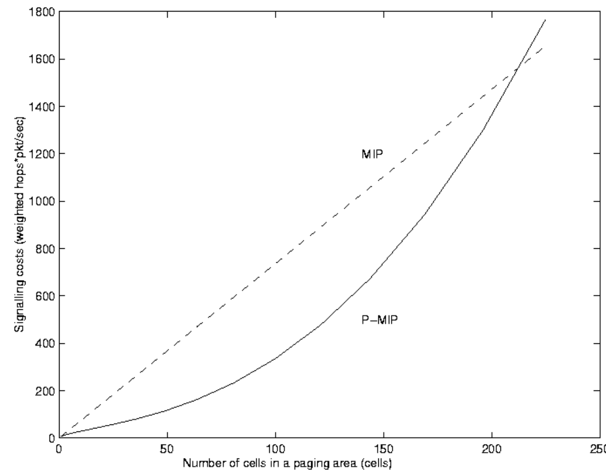


**Figure 3: Effect of Paging Area Size on Signaling Cost for Macro Cellular Systems**

As shown in Figure 3, the signaling cost for Mobile IP has a linear relationship with the number of cells. The reasoning behind this is that the more cells considered, the more cell boundary crossings, which causes more registration signaling. When paging is introduced, signaling includes paging and registration. We observe that when the number of cells in the paging area is under certain value, the signaling generated by paging remains small in comparison to the registration signaling overhead. Therefore, the total signaling cost is reduced more than increased by the addition of paging. When the paging area has a large number of cells, however, the signaling cost introduced by paging grows quickly. This is due to the fact paging request messages are unicast to all foreign agents in every cell in the paging area whenever paging is initiated. A number of observations can be made regards paging area size. The benefit of paging is limited when there are a small number of cells in a paging area. This is because of the large number of unnecessary registrations caused by boundary crossings. In contrast, the paging cost grows very quickly when there are a large number of cells in a paging area. As a result, this increases the total signaling cost rather than reducing it. Therefore, in order to optimize the total signaling cost and the system performance, it is important to select a suitable paging area size.

### 5.3.2 Hop Weight Ratio ($R_w$)

We set the number of cells in a paging area to 49 while other parameters remain the same as discussed above. The hop weight ratio ($R_w$) shows the different effects of paging and registration on signaling cost. When the ratio is small, local signaling (i.e., paging signaling) and core network signaling (i.e., registration signaling) have a similar impact on signaling cost. It is necessary to limit both paging and registration signaling in order to reduce the total signaling cost. As the hop weight ratio increases, the effect of paging signaling on signaling cost reduces, while the effect of registration signaling increases. Therefore, it is important to reduce the number of registrations to minimize the total signaling cost. In this case, an increase in paging can reduce the registration signaling. When the ratio increases to certain level, registration signaling dominates the total signaling cost, and the paging signaling cost becomes negligible.

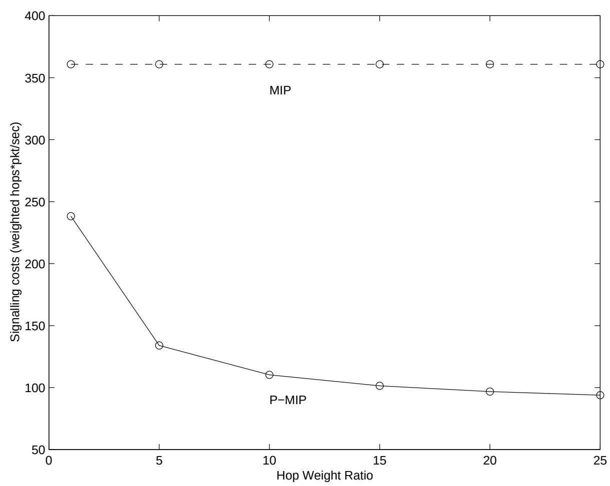


**Figure 4: Effect of Hop Weight Ratio on Signaling Cost for the Macro Cellular Configuration**

Figure 4 shows the signaling cost associated with P-MIP decreases as the hop weight ratio increases, while in contrast, the signaling cost with Mobile IP remains unchanged. Only registration signaling is used by Mobile IP where the weight ratio does not impact the total normalized signaling cost, see Equation (5). In contrast, paging and registration are used by P-MIP where the weight ratio has impact on the total signaling cost.

### 5.3.3 *Mobile Node Speed.*

Slow moving mobile nodes remain in the same cell for relatively long periods of time. As a result, the registration signaling due to boundary crossing is small in comparison. In this case, paging represents the major overhead because mobile nodes are likely to be located in the same network. But for fast moving mobile nodes, there is a large amount of registration signaling generated by cell boundary crossings. Paging brings significant benefits by reducing the number of registrations under these conditions.

### 5.3.4 *Mobile Density.*

In P-MIP, only a small percentage of mobile nodes are active, most nodes are idle. As discussed earlier, active mobile nodes in P-MIP are operationally equivalent to mobile nodes in MIP. The signaling cost of idle mobile nodes is relatively small in comparison to active mobile nodes. In P-MIP, an increase in mobile density only has impact on a smaller percentage of active users. In comparison, this affects all mobile nodes for MIP. As a result, signaling cost due to increased mobile density is more pronounced with MIP.

### 5.3.5 *Operational State.*

Because Mobile IP does not distinguish between active and idle mobile nodes, signaling cost is unaffected by the number of active mobile nodes. In P-MIP, however, the signaling cost increases as the percentage of active mobile node increases. In the extreme case where all mobile nodes are active, the signaling cost of P-MIP and Mobile IP are equivalent. A estimate for the percentage of active nodes in cellular telephony is between 10 and 20% [15].

### 5.3.6 *Data Session Rate.*

The data session rate has an impact of paging performance. However, the signaling cost is not affected by data session rate in MIP. In P-MIP, the entire paging area is flooded with a paging request message when a new incoming data session arrives. As the arrival rate of new sessions increases so does the total signaling. For lower session arrival rates, P-MIP reduces the registration signaling with minimal increase in paging overhead. The effect of the outgoing data session rate is similar to that of the incoming session data rate, except that the incoming rate has more impact on the signaling cost. This is because paging is associated with the incoming rate and not the outgoing rate.

## 5.4 Simulation Environment

We now study the performance of P-MIP using our ns simulation environment. This allows us to investigate scenarios that cannot be considered by analysis, as well as to validate our analytical results.

Constant bit rate (CBR) sources are used as traffic sources. The data session rate represents the number of times constant bit rate packet trains are sent to a mobile node during the simulation. Each data session lasts 10 seconds so that the active mobile node percentage is approximately 5%. The duration of each simulation experiment is 240 seconds.

The number of cells in a paging area is 4, 9, 16, 25, 36, and 49, respectively, where the paging area is square-shaped. The maximum velocity is set to 10 m/s and 5.2 m/s, respectively.

The data session rate for each mobile node is set to one in the paging area size experiment; that is, one constant bit rate packet train is destined to each mobile node for the duration of each simulation. Each session is operational for 10 seconds which is related to active mobile node percentage discussed in Section 5.3.

## 5.5 Simulation Results

We record registration and paging signaling for all experiments where the signaling cost is calculated based on these signaling counts. As discussed in Section 5.3, the registration refresh signaling is not considered. In order to validate the analysis, we use the same assumptions and system configuration as discussed earlier whenever possible. The average numbers of hops is $d_{FA,HA} = 16$ $(hops)$, the ratio is $d_{FA,FA} = \sqrt{n}$ $(hops)$, hop weight ratio $R_w = 8$. Half the hops between home agent and foreign agents are in the IP core network and half are in the wireless access networks, where $R_{local} = R_{core} = 0.5$.

### 5.5.1 Paging Area Size.

All mobile nodes operate with a pause time of 0, supporting continuous motion. This movement pattern has some similarity to the fluid flow model used in the analysis in which all mobile nodes move constantly. Some differences exist, however. First, speed and direction are continuously changing for each mobile node under simulation but are constant in the fluid flow model. Next, in the simulation environment, mobile nodes are confined to operate within two paging areas. In contrast, mobile nodes are moving into the paging area in the fluid flow model while other nodes are moving out.

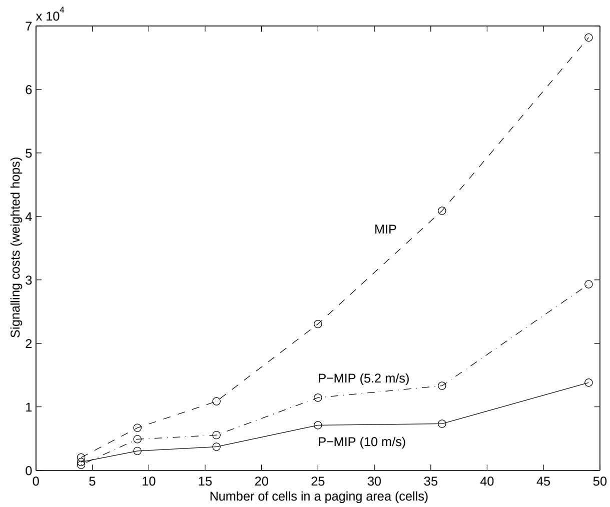


**Figure 5: Effect of Paging Area Size on Signaling Cost**

Simulation results for maximum velocity 10 m/s and 5.2 m/s are shown in Figure 5. The simulation results follow the same trends shown in Figure 3. Differences are likely to be associated with the different mobility models used. Simulation results confirm that in order to improve system performance using paging, it is important to select the appropriate number of cells in a paging area. Results confirm that paging is ineffective when there are too few or too many cells in a paging area.

We also observe from Figure 5 that paging provides more gain under fast mobility conditions. The reason for this is intuitive. The faster the mobiles move, the more boundary crossings experienced and registration signaling generated. Thus, paging can reduce signaling cost under higher mobility conditions.

We also investigate the impact of the discrete movement of mobile nodes that cannot be observed by using the fluid flow model used in the analysis. All mobile nodes are now set to operate with a pause time of 60, supporting discrete motion, in contrast to the continuous movement discussed above. Mobile nodes move to a destination, stay there for certain period of time and then move again. The results are similar to those with fast and slow moving mobile nodes with similar reasoning.

### 5.5.2 Data Session Rate.

In this experiment, we consider the effect of the data session rate on the total signaling cost. Simulation and analysis consider an active mobile node percentage to be approximately 5% [6]. This is configured as follows. For analysis, we explicitly specify the parameter $\alpha$. Under simulation, we obtain this implicitly by setting the data session holding time. For this experiment, the pause time is set to 0 and paging area size to 49.

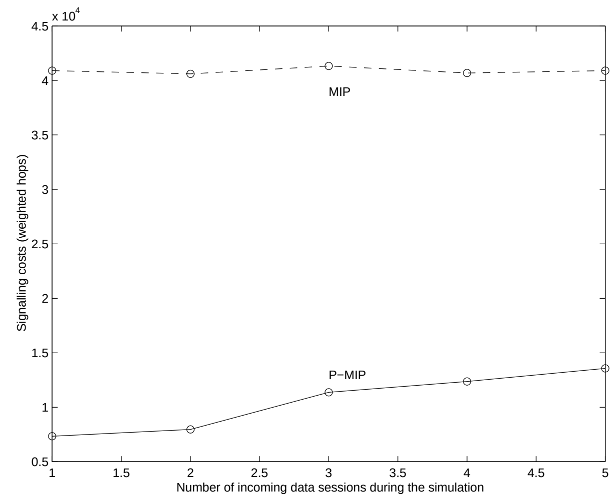


**Figure 6: Effect of Data Session Rate on Signaling Cost**

We observe that the simulation results, shown by Figure 6, follow the same trends as the analysis results discussed in Section 4.

From the simulation, we conclude that paging reduces the signaling cost significantly when the data session rate is low. In this case, paging reduces unnecessary registration without introducing paging overhead. More paging is required at higher data rates, and as a result, paging increases the signaling cost. Paging eventually degrades the system performance at very high data session rates. Signaling cost in Mobile IP does not change with the data session rate, as indicated in Section 4 and Figure 6.

## 6. CONCLUSION

In this paper, we have introduced a set of simple paging extensions for Mobile IP protocol called P-MIP. While P-MIP extends the base protocol it is also backward compatible with Mobile IP. Therefore, Internet Service Providers have the freedom to configure the P-MIP paging capability on an as needed basis without affecting the operation of the entire Mobile IP enabled Internet network.

Both the analysis and simulation results presented in this paper show that by carefully selecting suitable system configuration parameters (e.g., paging area size, active timer, etc.) paging can reduce unnecessary registration thereby reducing the overall signaling cost with the result of improved system performance. Paging also minimizes the impact of mobility rate, cell size and mobile node density on the signaling cost.

## 7. ACKNOWLEDGEMENT

The National Science Foundation (NSF) under the wireless technology award AW1-997943 has supported this research.

## 8. REFERENCES


[1] C. Perkins, "IP Mobility Support," RFC 2002, IETF, Oct. 1996

[2] X. Zhang, J. Castellanos, and A. Campbell, "P-MIP:Minimal Paging Extensions for Mobile IP," Internet Draft, IETF, Jul. 2000.

[3] Seiemund H. Redl, Matthias K. Weber, and Malcolm W. Oliphant, *An Introduction to GSM*, Artech House Publishers, MA, 1995.

[4] Gunnar Heine, *GSM Networks: Protocols, Terminology, and Implementation*, Artech House Publishers, MA, 1998.

[5] Michael D. Gallagher and Randall A. Snyder, *Mobile Telecommunications Networking with IS-41* McGraw-Hill, NY, 1997.

[6] A. Campbell, J. Gomez, C-Y. Wan, Z. Turanyi, and A. Valko, "Cellular IP," Internet Draft, IETF, 1999

[7] R. Ramjee, T. La Porta, S. Thuel, K. varadhan, and L. Salgarelli, "IP Micro-mobility support using HAWAII," Internet Draft, IETF, Dec. 1999

[8] R. Ramjee, T. La Porta, and L. Li, "Paging support for IP mobility using HAWAII," Internet Draft, IETF, Dec. 1999

[9] E. Gustafsson, A. Jonsson, and C.E. Perkins, "Mobile IP Regional Registration," Internet Draft, IETF, Mar. 2000

[10] H. Haverinen and J. Malinen, "Mobile IP Regional Paging," Internet Draft, IETF, June 2000

[11] C. Castelluccia, "Extending Mobile IP with Adaptive Individual Paging: A Performance Analysis," INRIA, Nov. 1999 *http://www.inrialpes.fr/planete/people/ccastel/ Welcome.html#PUB*

[12] T. Brown and S. Mohan, "Mobility Management for Personal Communications Systems," *IEEE Transactions on Vehicular Technology*, vol. 46, No. 2, pp.269–278, May 1997

[13] S. Mohan and R. Jain, "Two User Location Strategies for Personal Communications Services," *IEEE Personal Communications*, pp.42–50, First Quarter 1994

[14] R. Ramjee, T. La Porta, S. Thuel, K. varadhan, and S.Y. Wang, "HAWAII:A Domain-based Approach for Supporting Mobility in Wide-area Wireless Networks," *http://www.bell-labs.com/user/ramjee/papers/papers.html*

[15] A. G. Valko, *Personal communication with A. G. Valko*, Sept. 2000

[16] J. Widmer, "Network Simulations for A Mobile Network Architecture of Vehicles," *http://www.icsi.berkeley.edu/ widmer/mnav/ns-extension*

[17] K. Fall and K. Varadhan, editors. "ns Notes and Documentation," *http://www-mash.cs.berkeley.edu/ns*

[18] J. Broch, D. Maltz, D. Johnson, Y-C Hu, and J. Jetcheva, "A Performance Comparison of Multi-Hop Ad Hoc Network Routing Protocols," In Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 98), ACM, Dallas, TX, Oct. 1998.