Caching Location Data in Mobile Networking

Shyhtsun F. Wu *

Computer Science Department
Columbia University
New York, NY 10027
wu@cs.columbia.edu

Charles Perkins

Mobile Systems Design
IBM T. J. Watson Research Center
Yorktown Heights, NY 10598
perk@watson.ibm.com

Pravin Bhagwat †

Computer Science Department
University of Maryland
College Park, MD 20742
pravin@cs.umd.edu

Abstract

Each *Location Directory (LD)* provides the information to locate the forwarding address of a *Mobile Host (MH)* in a mobile networking system. The LD should be updated when the MH moves from one access zone to another access zone. If an *Internet Access Point (IAP)* (or base station) could cache some useful LD information entries, then, after receiving a packet from an MH in its access zone, the IAP could check its cached LD and make an optimal routing decision. With the LD distributed over the Internet Access Points, we avoid *Dogleg Routing* because we no longer need to forward the packet to the *Mobile Router (MR)* where a master copy of the LD is located. In this paper, we investigate how the LD data structure is used, updated, and distributed among Mobile Routers, Internet Access Points, and Mobile Hosts. Furthermore, we present our LD design and implementation in the mobile IP network.

1 Introduction

Recently, the problem of providing continuous network connectivity to mobile computers has received considerable attention [8, 9, 2, 1, 4, 6, 7, 3]. One key issue in achieving high performance mobile networking is to provide optimal routing with dynamic location directory (LD) information. Under the assumption that mobile hosts in a mobile networking system move randomly and frequently, it is critical to have an efficient scheme to distribute the dynamic LD information over routers, access points, and mobile hosts.

In this paper, we present a scheme to efficiently distribute location directory information. Our scheme works with both of the most commonly used mobile redirection schemes, loose source routing (LSR) and encapsulation (ENC) (see [4]). In the next section, we give some background information about the mobile networking system. Then, we discuss why it is important to distribute location directory information over the IAPs (internet access points). Finally, our scheme as well as its implementation is presented.

2 Background

Our system involves the participation of three types of entities - viz., *Mobile Host (MH)*, *Internet Access Point (IAP)* and *Mobile Router (MR)*. The networking architecture that we assume is that of a set of IAPs connected through a wired backbone. An IAP supports at least one wireless interface and functions as a gateway between the wired and wireless side of the network. Due to the limited range of wireless transceivers, a mobile host can set up a direct link layer connection with an IAP only within a limited geographical region around it. This region is referred to as an IAP's *access zone*. The geograph-

*Wu is supported by an IBM Fellowship.

†Bhagwat is supported by an IBM Fellowship.

ical area covered by an access zone is a function of the medium used for wireless communication. The range of infrared access zones is typically limited to about 20 feet, while that of radio frequency zones could be significantly larger.

Within one campus or administrative domain there could be multiple (sub)networks reserved for mobile hosts. Each (sub)network has a router which is referred to as Mobile Router (MR). Unlike other routers, an MR is not required to have an interface corresponding to the wireless (sub)net it serves. If an MR has a wireless interface then it can also function as an IAP. The association between an MH and its current IAP is kept in the location directory (LD), which is maintained at the MR.

A mobile host retains its address regardless of which IAP's access zone it is in. It can start sessions with other hosts (both mobile and stationary) and move into other IAP's access zones without disrupting any active sessions. The movement of a mobile host is completely transparent to the running applications, except possibly for a momentary pause which may occur while the access zone switch takes place. An MH can reside in the zone of only one IAP at any given time. Even if zones of two IAPs spatially overlap, an MH routes its outgoing packets through only one of them. An IAP can have multiple MHs in its access zone. We use the term *Correspondent Host (CH)* to refer to the host communicating with an MH.

3 Replicating the Location Directory (LD)

We assume that the location directory (LD) information of a mobile host, MH_{dst} is kept in only one mobile router, MR_{foo}. That is, no other mobile router can have the LD information about MH_{dst}. When MH_{dst} moves into a new access zone (step 1 in Figure 1), IAP_{dst}, it sends MR_{foo} a special message containing the new location information through IAP_{dst} (steps 2 and 3). After receiving this special message, MR_{foo} updates the location directory and assigns a new sequence number to this entry (step 4). Therefore, later packets (steps 7 and 8) destinated for MH_{dst} would be forwarded to IAP_{dst} by MR_{foo} (steps 8 and 9).

We have another mobile host MH_{src}, attached to IAP_{src}, which would like to talk to MH_{dst}. And, MR_{foo} might not have location information for MH_{src}; MH_{src} belongs to MR_{bar}, and MR_{foo} advertises reachability for the subnet of MH_{src}. The packets from MH_{src} to MH_{dst} would always pass through MR_{foo} (steps 1 and 2 in Figure 2) because only MR_{foo} has location information for MH_{dst}. Even if IAP_{src} and IAP_{dst} are very close to each other, the packets still need to take a sub-optimal route to access MH_{dst}. This is the *DogLeg Routing* (or *Triangle Routing*) problem mentioned in [5]. The solution is to let IAP_{src} cache the location information about MH_{dst} (step 6) so that the packets follow an optimal route from MH_{src} to MH_{dst} (steps 7, 8, 9, and 10 in Figure 2).

4 Distributing the LD Information

In the previous section, we suggested that distributing LD information can improve the mobile routing performance. In this section, we consider how to efficiently distribute the location directory information over the Internet Access Points. Our first design goal for distributing LD information to IAPs is that the LD entries should not be distributed to a IAP unless the IAP would almost definitely use these entries. Our second goal is to avoid or to minimize the extra processing overhead on every incoming packet.

4.1 Protocol

4.1.1 Location Directory Information Update Message (LUM)

A *Location directory information Update Message (LUM)* is a location update entry for a mobile host. It contains four fields: the address of the Mobile Host (Destination), the forwarding address for this Mobile Host (Gateway), a unique sequence number incremented by the Mobile Host (MH_{dst}), and a special flag to identify the source of this LUM. We use this special flag (SF) to identify whether the LUM message is for IAP_{src} or IAP_{dst}.

4.1.2 Mobile Router (MR)

1. If the received packet is destined to this MR (step 2 in Figure 3), then pass it to the upper layer. Otherwise, go to the next statement.
2. Check whether the destination is a mobile host belonging to this MR. (This is done by checking the route flag in the MH's route entry.) If yes, it means that the source, MH_{src}, of the packet does not know the current location of the destination MH (MH_{dst}).
3. Forward the packet (step 3 in Figure 3).
4. Send a LUM to MH_{src}, the CH of MH_{dst} (step 5 in Figure 3), about MH_{dst}.

4.1.3 Internet Access Point

1. Receive a LUM from a MH in its access zone (steps 7 and 11 in Figure 3). Or, snoop a LUM from MR to the MH.
2. Check if a corresponding location cache exists (step 8 and 12). If so, we compare the sequence numbers to decide whether we will update this entry. (This is only necessary if we snoop packets here.)
3. If the LUM entry is updated and the SF flag in the LUM message is on, then we turn off that SF and send the LUM to MH_{dst} (step 9 in Figure 3). This is done so that IAP_{dst} will not send IAP_{src} a LUM.

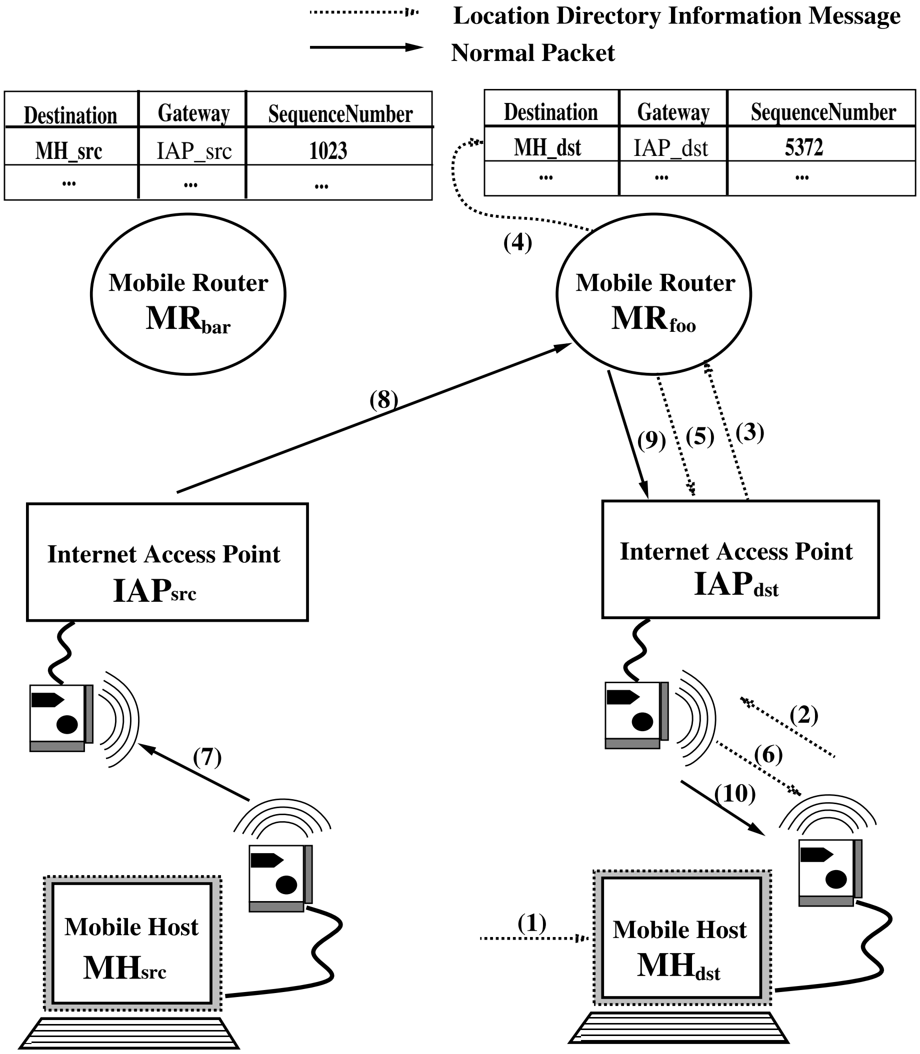

Figure 1: Mobile Networking without Distributing LD Information over IAP's

4.1.4 Mobile Host

1. Receive an LUM from either the MR (steps 5 and 6 in Figure 3) or IAP of its correspondent host (steps 9 and 10 in Figure 3).

2. Check whether the LUM entry exists, If yes, we compare the sequence numbers to decide whether we will update this entry.

3. Send the same LUM to its current IAP (steps 7 and 11 in Figure 3).

4.2 Example

When MR_{foo} needs to forward a packet from MH_{src} to IAP_{dst}, it knows that IAP_{src} does not know where IAP_{dst} is. If this were not true, then this packet would never have arrived at MR_{foo}. Then, MR_{foo} should distribute the location information to IAP_{src} since this forwarding information is useful for IAP_{src}. Furthermore, in a lot of applications, the communication is bi-directional, and therefore, MR_{foo} should also inform the IAP_{dst} about the current location of MH_{src}. However, only MR_{bar} has the location information about MH_{src} and MR_{foo} in this case will not know whether MH_{src} is a mobile host or a stationary host. So, on one hand, MR_{foo} can not send a message to IAP_{src} because the address of IAP_{src} is unknown. On the other hand, MR_{foo} can not inform IAP_{dst} about the current location of MH_{src} for the same reason.

Our solution is the following: MR_{foo} distributes the location information about MH_{dst} directly to MH_{src}. And, since MH_{src} has the knowledge about IAP_{src}, it forwards the information to its current IAP_{src}. Then, after knowing the relationship between MH_{dst} and IAP_{dst} from MH_{src}, IAP_{src} sends

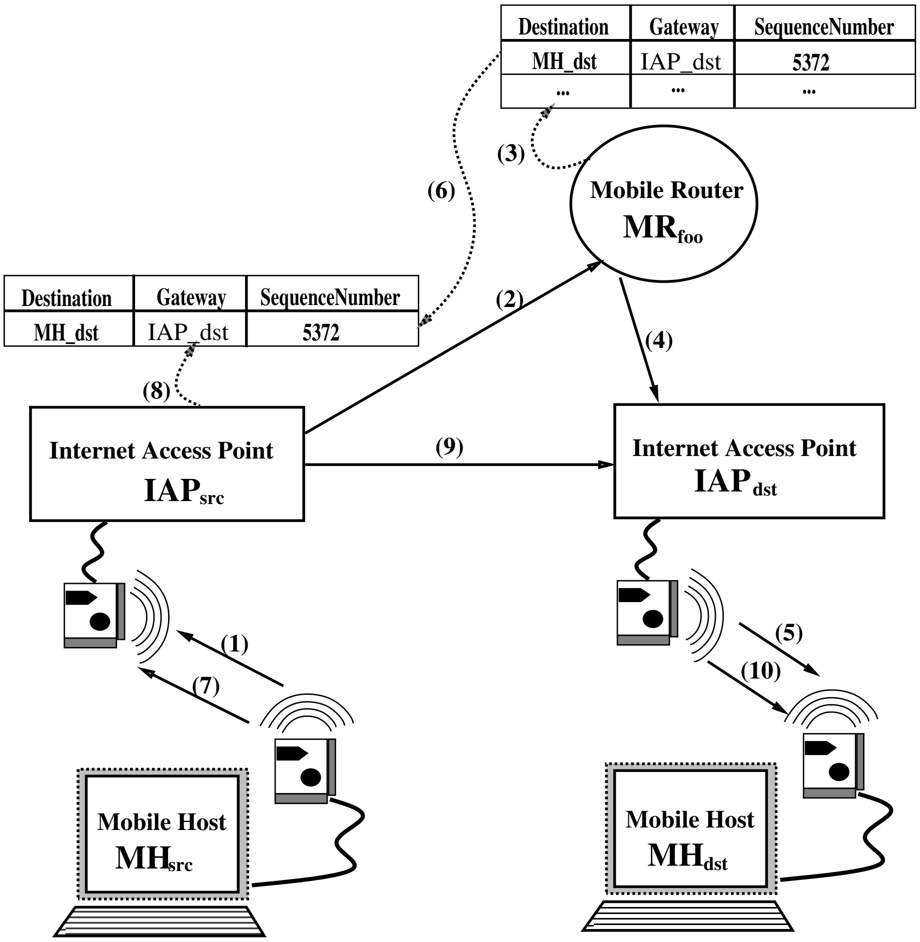

Figure 2: Distributing Location Directory Information over IAP's

the location information about MH_{src} to MH_{dst}. And, again, MH_{dst} would forward this information to IAP_{dst}. After receiving this message from MH_{dst}, IAP_{dst} would NOT send another LUM to MH_{src} because the SF flag is off and therefore it knows IAP_{src} has already received the LUM from MR_{foo}. Finally, as depicted in Figure 3, both IAP_{src} and IAP_{dst} have enough information to route the packets between MH_{src} and MH_{dst} without passing through any mobile router.

MH_{src} may receive LUMs for MH_{dst}, not only from MR_{foo}, but also from IAP_{dst}. Before the IAP_{dst} receives the LUM from IAP_{src}, MH_{dst} sends a packet back to MH_{src}. This packet arrives at MR_{bar}, and similarly, MR_{bar} sends MH_{dst} a LD message, which is then forwarded to IAP_{dst}. At this point, IAP_{dst} does not know whether IAP_{src} has received the LD message from MR_{foo}, and therefore, sends a LD message to MH_{src}. MH_{src} can determine which message is the most recent one as long as each message contains a sequence number generated by MR_{foo} (originally from the time stamp information sent by MH_{dst} when MH_{dst} last detected an access zone switch.)

5 Mobility of Cached Location Directory Information

One important issue in mobile computing is to provide continuous network service for hosts as they move around. In our example, both MH_{src} and MH_{dst} can move into new access zones. Therefore, the distributed location directory information on IAP_{src} and IAP_{dst} might need to be updated. Furthermore, the new IAPs (*i.e.,* IAP_{src}^{new} and IAP_{dst}^{new}) for both mobile hosts (*i.e.,* MH_{src} and MH_{dst}) also need the location directory information for optimal mobile routing.

Whenever MH_{dst} receives an LUM, it will keep it in an LD table. Then, when MH_{dst} moves to another access zone IAP_{dst}^{new}, it can send its cached LD information entries with sequence numbers to the IAP_{dst}^{new}. So, immediately, the connections from MH_{dst} to other correspondent hosts (CHs) would still follow the optimal route. However, this might not be true for all connections. For example, if

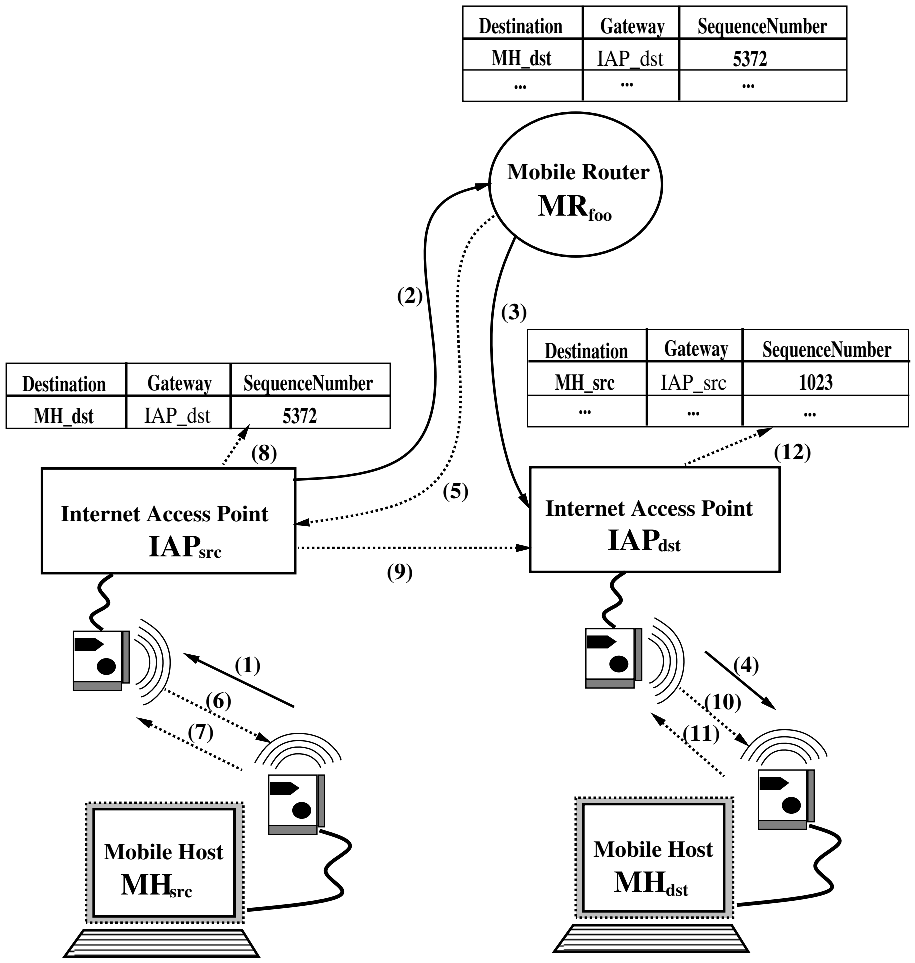

Figure 3: An Example of Distributing Location Directory Information

another mobile host $MH_{another}$ in IAP_{dst} tries to communicate with MH_{src} before MH_{dst} does. And, the LD entry for MH_{src} was built before MH_{dst} used it. Then, a sub-optimal route would be followed. In this case, the connection between MH_{src} and MH_{dst} would pass through neither MR_{foo} nor MR_{bar}. Thus, MH_{dst} would not have the LD entry in its local LD copy. After moving to a new access zone, MH_{dst} might need to go over the LD distributing process again to get the optimal route. This results from our policy of tagging each location cache entry with just one MH, not multiple MHs, even when the cache entry is used by multiple MHs.

Another problem is that, after the movement of MH_{dst}, the LD information kept by both MH_{src} and IAP_{src} is invalid. The first packet will still go to IAP_{dst}. Then, the packet will be forward to MR_{foo} and the whole LD distributing process starts again. In the case that IAP_{dst} is down, we can invalidate the cached LD entry upon receiving an ICMP host/network unreachable message for IAP_{dst}. We call this a *lazy* invalidating approach.

An alternative approach is to *eagerly* invalidate the LD entries on all the IAP's having an MH_{dst} LD entry immediately after MH_{dst} moves to a new access zone. This approach taken by [8, 9] has two major shortcomings:

1. The new updated entries might not be used before the next entry update. Thus, we waste some network bandwidth.

2. It is expensive to track which internet access points have an entry for MH_{dst}. In our implementation, it requires one extra route table lookup and possibly the manipulation of a linked list data structure for EACH incoming encapsulated or loose source routed IP packet.

Thus, to avoid unnecessary and extra processing, we do not take this eager approach.

6 Implementation and Evaluation

We have implemented our scheme on a set of IBM PS/2 model 80 running AIX version 1.2. To make the LD distribution scheme work, we only needed to modify the AIX kernel on the mobile router. We use the same kernel as described in [1] with an extension to support both LSR and encapsulation for IAP and MH. We built three user level processes: *mh_serv*, *iap_serv*, and *mr_serv* for mobile host, internet access point, and mobile router respectively, to distribute the LD information. All user processes communicate with each other via UDP sockets. Furthermore, the cached LD entries with sequence numbers are kept in the user processes. And, the first two columns of each entry (*i.e.,* omitting sequence numbers) are replicated into the kernel routing table.

For better performance, we made another modification for the IAP kernel. We introduce a snooping mechanism in the kernel of IAP to grab the LD messages for all the MHs in its access zone. By doing this, IAP can immediately have the LD information without waiting for the round-trip UDP message from one of its MHs. Although this means that IAP kernel has to check every incoming packet, the checking is very efficient. And, only on the IAP kernel, it requires two extra comparisons: the UDP protocol number and the *mh_serv* special port number in the *ip_forward* function. Thus, the IAP can immediately send another LUM to the IAP of the corresponding mobile host.

If IAP does not have the snooping capability, the LD information will still arrive but with some extra delay. On the other hand, if MH can not handle the LUMs correctly, the snooping mechanism would still be able to install the LD entries for optimal mobile routing. However, in this case, when this MH moves, it will not be able to provide the LD entries for the new IAP. Finally, if both MH and IAP do not work as we expect, then packets from the MH would still arrive at the correct destinations through some sub-optimal routes.

7 Remarks

We present an efficient way to cache the location data information among all the components in a mobile networking system. The scheme is simple and few changes need to be made to the kernel. The required network bandwidth is small because only useful location directory information will be sent. Furthermore, our scheme avoids extra processing overhead for each packet.

The problem in general can be treated as a cache coherence control problem. However, we observed the required consistency in distributed LD is much weaker than the consistency criteria in a distributed shared memory system. It would be interesting to know the LD consistency criteria in a formal way, and thus we can compare and analyze different LD distributing schemes, *e.g.,* [2, 8, 9]. Furthermore, we need to investigate the relationship between the network performance and the consistency criteria.

References

[1] Pravin Bhagwat and Charles Perkins. A Mobile Networking System based on Internet Protocol (IP). In *Usenix Symposium on Mobile and Location-Independent Computing*, August 1993.

[2] John Ioannidis, Dan Duchamp, and Gerald Q. Maguire Jr. IP-based Protocols for Mobile Internetworking. In *Proceedings of ACM SIGCOMM*, pages 235–245, 1991.

[3] David B. Johnson. Transparent Internet Routing for IP Mobile Hosts. Internet draft, July 1993.

[4] Andrew Myles and David Skellern. Comparing Four IP Based Mobile Host Protocols. In *Joint European Networking Conference*, May 1993.

[5] Charles Perkins. Providing Continuous Network Access to Mobile Hosts Using TCP/IP. In *Joint European Networking Conference*, May 1993.

[6] Charles Perkins and Yakov Rekhter. Short-cut Routing for Mobile Hosts. Internet draft, July 1992.

[7] Charles Perkins and Yakov Rekhter. Support for Mobility with Connectionless Network Layer Protocols. Internet draft, November 1992.

[8] Fumio Teraoka, Yasuhiko Yokote, and Mario Tokoro. A Network Architecture Providing Host Migration Transparency. In *Proceeding of ACM SIGCOMM*, Sept 1991.

[9] Hiromi Wada, Takashi Yozawa, Tatsuya Ohnishi, and Yasunori Tanaka. Mobile Computing Environment Based on Internet Packet Forwarding. In *proceeding of Winter USENIX*, pages 503–517, San Diego, CA, Jan 1993.