Mobility Management Meets Privacy–the Failure of Existing Proposals and a New, Future-Proof Approach (Extended Version)

Christian Hauser

04.08.2004
20 Seiten / Pages

Privacy, Mobility Management

Identification of threats to the approach of multiple identities resulting from mobility management. Evaluation of existing proposals regarding these threats. Proposal of a new approach providing better protection.

Protection of user privacy will gain increasing importance in future mobile systems. This is due to the trend to context-awareness as well as to open systems in terms of publicly available interfaces and of an open provider model. To meet the users' privacy needs the approach to allow them to act under several different pseudonyms or personas is promising. Crucial attacks on this approach are the linking of different pseudonyms of a user and the inference of knowledge that the user wants to conceal in the context of a pseudonym. Both attacks can be executed with application knowledge as well as with information originating from the communication process itself. Here, mobility management plays a central role as it reflects the user’s behavior. Thus, privacy protection by multiple pseudonyms can only be successful if the communication system, in particular mobility management, is privacy-aware. In this paper, the threats to the approach of multiple pseudonyms resulting from IP-based mobility management is analyzed. Existing proposals are evaluated and it is shown, that they do not offer sufficient protection. At the end, an outlook to a new approach to mobility management protecting multiple pseudonyms is given.
Mobility Management Meets Privacy--the Failure of Existing Proposals and a New, Future-Proof Approach (Extended Version)

Christian Hauser
Institute of Communication Networks and Computer Engineering
University of Stuttgart
Pfaffenwaldring 47
70569 Stuttgart
Germany
hauser@ikr.uni-stuttgart.de

Abstract
Protection of user privacy will gain increasing importance in future mobile systems. This is due to the trend to context-awareness as well as to open systems in terms of publicly available interfaces and of an open provider model. To meet the users’ privacy needs the approach to allow them to act under several different pseudonyms or personas is promising. Crucial attacks on this approach are the linking of different pseudonyms of a user and the inference of knowledge that the user wants to conceal in the context of a pseudonym. Both attacks can be executed with application knowledge as well as with information originating from the communication process itself. Here, mobility management plays a central role as it reflects the user’s behavior. Thus, privacy protection by multiple pseudonyms can only be successful if the communication system, in particular mobility management, is privacy-aware. In this paper, the threats to the approach of multiple pseudonyms resulting from IP-based mobility management is analyzed. Existing proposals are evaluated and it is shown, that they do not offer sufficient protection. At the end, an outlook to a new approach to mobility management protecting multiple pseudonyms is given.

1 Introduction

Today’s Situation
In today’s mobile communication systems clear roles are assigned to the participating peers. On the one hand, there are untrusted consumers and on the other hand there are large operators as service providers. Communication services are mostly provided by a single network operator which may also provide value added services, e.g., location-based services. Even wireless high bandwidth hotspots are mostly operated by those large providers, today. This single provider is well-known and an abuse of personal user information would cause a significant loss of credibility. Therefore, most private users trust their provider regarding protection of private information, today.

Trends in Communication Systems
A wide-spread trend in communication systems is on open philosophy like in the Internet. Beyond publicly available interface specifications, this also means—in its final consequence—that everybody may participate in the system not only as an information consumer but also as a content or service provider. This trend is explicitly supported by today’s incumbent operators as a large variety of services also increases the revenue of the network operator. Thus, it can be assumed, that in future there will be many service providers for certain services.
Considering recent developments in networking, e.g., the wide spreading of IEEE 802.11 networks and the huge amount of addresses available in IPv6 [1], virtually everybody can act even as a communication service provider by either offering the network for roaming users or by providing a Mobile IP Home Agent and Home Addresses for mobile users. Like today there are many free providers for E-mail communication, future can be assumed to bring many mobility management providers for IP-based communication. As the users will not always know these providers and they are not always large well-established companies, this situation can lead to varying or even diminishing trust of users in providers.

Moreover, private users—who want to protect their privacy—will be required to be reachable by the communication system. This is due to the loosening of roles of communication system participants, i.e., due to services on private persons’ devices no longer being assumed only as information consumers, i.e., clients, but also as providers, i.e., servers. An example for that can be a service for the provision of traffic data or for the user’s preferences for (virtually) meeting interested persons. This is also necessary in new communication services like, e.g., event services in which a user subscribes to an event and will be notified when this event occurs later in time. This requirement in principle is already solved and deployed from a communications perspective. In contrast, it is often not considered by existing privacy work, e.g., [2], [3], [4], [5], [6].

Regarding the access network infrastructure the trend is towards a 4th generation mobile communication infrastructure. This means mainly an integration of all possible access technologies, like UMTS, IEEE 802.11 WLAN or Ethernet. Therein, IP is going to be the common communication platform and mobility management is going to be handled on network layer by Mobile IPv6¹ [7]. Research already studies so-called all-IP approaches with mobility management being done on IP layer only, e.g., [8]. This will be considered in this paper to analyze the resulting threats.

A further evolution is that location information maintained by mobility management is assumed to become more accurate due to networks with a smaller geographical extent. First this is due to the wide spreading of IEEE 802.11 WLANs and secondly due to smaller cells in next generation cellular systems, e.g., UMTS, compared to 2nd generation systems like GSM.

An additional common trend in communication systems going beyond mobility support is context-awareness. In these systems, the user’s personal context is exploited to provide additional and personalized services. Examples for context information can be the user’s location, capabilities of the communication device or the user’s current situation. Consequently, a lot of personal information of the user is disclosed to the system. This goes far beyond the information known by today’s mobile communication systems which only know very roughly the user’s location for paging purposes, the billing address etc.

An exact justification of these trends is out of the scope of this paper. Details can be found, e.g., in [8], [9] and [10].

Following these arguments, the traditional trust model of mobile communication systems, in which the users fully trust their single provider, is no longer suitable for future systems. Thus, new protection mechanisms have to be adopted. Therein, service and network providers must be considered as potential attackers.

¹ If not explicitly stated otherwise, the term Mobile IP refers to Mobile IPv6 throughout this paper.
**Protection Approach**

According to the right on informational self-determination the user shall be able to determine who may see which information about him. A promising approach to achieve this is the possibility to act under multiple pseudonyms or personas. Thereby, the user is able to separately tune the personal information revealed in the context of each pseudonym. E.g., there can be a pseudonym revealing the user’s location in the context of a navigation service and another pseudonym with the user’s name and address for M-business. In neither context it is necessary to reveal both, location and name. This approach allows the user to split and control the trace of personal data disclosed to communication partners.

Communication partners do not only see data the user explicitly releases on application level as being part of the pseudonym’s context. They also see additional information which is evolving in the context of service provisioning or the communication process itself. To emphasize this difference, the term virtual identity (VID) is used throughout this paper. A VID comprises the complete view a certain entity has on a user. This is in line with the common use of the term identity in reality, in which the user’s identity is defined not only by the name but by the entirety of the personal information, e.g., the behavior and the habits.

![Figure 1.1: Example of VIDs and their augmentation](image)

Figure 1.1 shows an example of a user acting under VID 1 towards a Navigation Service and under VID 2 towards an Online Shop. In this example, the user’s communication partners see application data—the pseudonyms, the name and the location trace—as well as information evolving from the communication process which is the IP address.

The crucial threat when using VIDs is that an attacker can infer additional information, thus augmenting its view on the user. In the example, this can happen first of all by collaboration of the Online Shop and the Navigation Service by utilizing the identical unique data being common to both VIDs. Here, this unique common data is the IP address. By the identical IP address, they can link the VIDs as belonging to one single user and thus, merge their knowledge about this user. In general, this common unique data may stem from the applications, e.g., the same credit card number, as well as from the communication system, e.g., the same IP address. Generally spoken, the crucial problem is the user’s unique behavior underlying all his VIDs. In the communication system, this unique behavior is reflected in the mobility management which has to keep track about the user’s movement in order to assure reachability. Therefore, mobility management plays a key role in this privacy protection approach.

The second possibility for an augmentation of an attacker’s view is inference of new information from known information. In the example, the Navigation Service can infer the sensitive information about the user by the location trace. This is possible due to the fact that a user’s
location often discloses more sensitive information, e.g., on the visit of a hospital. Again, mobility management in which the user’s location is stored plays a central role here.

The discussion of augmentation threats and their origins shows that a solution for protecting VIDs which focusses on application layer only is insufficient. It cannot prevent attacks based on information resulting from mobility management.

**Contribution**

This paper presents the first holistic evaluation of the augmentation threats of mobile IP-based communication to VIDs. Therein, not only the threat of inference of a pseudonymous user’s personal information, e.g., the activity, is considered but also the threat of linking several VIDs. The analysis considers a pseudonymous user in both roles, sender and recipient. Existing systems for mobility management and anonymous communication are classified and evaluated regarding these threats. Moreover, a new approach for a mobility management system considering all threats is proposed.

**Structure**

The rest of the paper is structured as follows. At first IP-based mobility in the context of VIDs is analyzed and the resulting threats are presented in chapter 2. In chapter 3, the protection of existing and proposed systems against these threats is evaluated. Chapter 4 proposes a new approach to solve the threats derived in the analysis. Finally, chapter 5 concludes and gives an outlook to future work.

## 2 Analysis of IP-based Mobility in the Context of Multiple VIDs

In this chapter, the threats to VIDs resulting from mobile IP-based communication are explored by analyzing the sensitive information maintained by the mobility management system. After an exploration of privacy implications of fixed IP communication in general, new impacts by VIDs are evaluated. Finally, the analysis is extended to the impact of mobility.

In the following threat analysis it is generally assumed that a user has a single device which has a single IP interface with one IP address. This is a simplification compared to real-life, in which a user can have several devices, e.g., a PDA as well as a mobile phone and a notebook. These devices can have several interfaces, e.g., a UMTS card as well as an IEEE 802.11 card with multiple IP addresses each. This simplification allows a clearer description of the problems without affecting the principle considerations and results. Moreover, it is assumed that IP addresses are globally routable and do not refer to the device, e.g., they do not have the link layer address encoded as may be done in IPv6 [11].

**Fixed IP Communication**

In packet-based communication there are two basic pieces of information necessary:

1. Each device connected to the network needs an identifier\(^1\) for addressing which is unique in the addressing domain. Each packet carries this identifier. It is used by the respective communication partners to address packets directed to the device as well as by the mobility management that needs to know which device is addressed. This identifier can in principle be chosen arbitrarily containing no information about the user. It only has to be unique in the network (or in the respective addressing subdomain).

---

1. In this paper, the terminology *identifier* and *locator* is adopted from the respective IETF discussions, e.g., in [12].
2. Each packet has a locator attached which indicates the location to where it has to be delivered, i.e., the topological location of the destination device. This locator does not necessarily have to be visible to the communication peers as they need not be aware of the topological location of their communication partner. The communication system has to see the locator as it is used for path finding.

In contrast to the identifier, the device’s locator is sensitive regarding privacy. First of all, the user is assumed to always have the device with him. Thus, the locator also denotes the topological location of the user. Moreover, in IP the topological location can be mapped to the geographical location with a certain accuracy. Knowledge of the user’s location, in turn, often allows inference of more sensitive information, e.g., the user’s current activity.

Furthermore, as the locator is identifying the user’s access network, it can be deduced which provider, which technology, etc. the user is using. It may even be possible to infer the user’s identity if only certain users can be at the respective (topological or geographical) location, e.g., when using a wireless LAN at a private place.

Concerning fixed IP-based communication, one specific property of IP is that the IP address serves as both, identifier and locator. This implies a crucial problem: The identifier can no longer be chosen arbitrarily and be anonymous but, as it is also the locator, it includes sensitive information. As the identifier must be known by the communication partners, the device’s—hence also the user’s—location is known to them and not only to the communication system. There exist approaches in the IETF to alleviate this problem, e.g., [12]. Nevertheless, these approaches are neither standardized yet, nor can be anticipated to replace or enhance IP in a conceivable timeframe.

Privacy Implications to VIDs

Considering VIDs, the problem is extended. When acting under multiple VIDs the user is visible in multiple contexts in the system as a whole, i.e., the communication system and application layer services. In the example above, this is one context for navigation and one context for M-business. These contexts shall not be linkable. On the other hand, there is only one device per user attached to the network. Therefore, at some place in the communication system, the VIDs contexts must be merged to a single context of the user’s device. This implies an inherent link between the VIDs that must be concealed against potential attackers. In practice, there exist several possibilities to merge the different contexts—sooner or later on the path a packet takes to the addressee’s device.

Figure 2.1 shows two intuitive possibilities for the merge. There are two boxes, indicating the knowledge of the Communication Partners and that of the Communication System. In general, a VID is addressed by an identifier which points to a locator which represents the device’s location. One possibility for the context merging is that all VIDs are addressed by the same identifier (black arrows). Another possibility is that VIDs have different identifiers but the user’s device only has one locator (grey arrows).

---

1. Without our simplification, the user can have multiple devices and (virtual) interfaces. Nevertheless, there will always be potentially more VIDs than devices and interfaces. If several virtual interfaces of the same physical interface will be used, they will all be in the same IP subnetwork and thus are also likely to be linked by an attacker. So the principle problem, that contexts are merged in the system remains.
Both realizations have different properties with regard to privacy as well as to scalability. In the first possibility, even communication partners can link VIDs as the identical identifier is visible to them. In the second possibility only the communication system provider can link VIDs as the common locator is only visible to the provider. On the other hand the first possibility requires only one state per user—the binding between identifier and locator—to be maintained by the system, whereas the second possibility would require one state per VID. Thus, the common identifier would be preferable from the scalability point of view. An exact evaluation of the trade-off with respect to scalability and performance is out of the scope of this paper, which analyzes the principle coherences with respect to privacy. Nevertheless, it must be stated that there will always be a trade-off between costs, e.g., in terms of scalability or performance and privacy. Therefore, the users should be in control of choosing maximum privacy or maximum performance.

In standard IP, in which the identifier is also the locator these two different possibilities of merging collapse into only one.

**Mobile Communication with VIDs**

Additional complexity is introduced by considering a dynamic system with mobile users. Regarding the identifier, there are no changes since it has to remain immutable in order to always allow addressing of the mobile user independent of the current location.

Regarding the locator there is a difference compared to a fixed system. The locator changes while the user’s device moves through different networks. Hence, correlation between the identifier and the locator becomes mutable and must be resolved by the mobility management at the time of packet delivery. This introduces time as a new dimension to be considered.

This dynamics of the locator implies the following new threats:

1. Linking of VIDs: The locator’s dynamic behavior implicates the possibility to link several VIDs via the identical mobility behavior, e.g., an identical trace or pattern of locator changes, in the context of the VIDs. User behavior is represented in the system not only by locators stored in the system but also by the point in time at which they are stored, erased, or changed.

2. Inference of sensitive information by several locators: The possibility of inference of more personal information than in the fixed scenario evolves as not only one locator is present but several locators are. Thus, it is, e.g., possible to track the user’s movement.

3. Inference of sensitive information by locator changes: Locator changes can be sensitive. This can be first due to the information of the move from a certain network to another specific network. Moreover, it can be due to the point in time when the change takes place, e.g., from the time of a change from a certain private WLAN to a public UMTS network it can be inferred when the tracked mobile user leaves the office.
4. Substantiation of imprecise knowledge: As the state of the system is changing, it is possible for an attacker to gain several different views on it. This can allow to substantiate a suspicion, e.g., about the user behind a VID, with methods like [13] from the area of artificial intelligence. This possibility gains on relevance when imprecise observation of sensitive information, i.e., suspicions, is considered.

Summing up the threats to VIDs in fixed and mobile IP-based communication, the structure depicted in Table 2.1 is evolving. There are two main threats: Linking of different VIDs of one user shown in the upper part and inference of personal information, which is shown in the lower part. The threats are named by abbreviations. The first part indicates whether it is a link threat (Link) or an inference threat (Inf). The second one indicates whether it is a threat also appearing in the fixed scenario (F) shown in the left part or only in the mobile scenario (M), which is shown in the right part. This can also be considered as the distinction whether a threat is also present in case of a short observation during which the state of the system does not change or only in case of a longer observation. Inference threats have a third letter indicating whether the sensitive information is inferred from the locator (L) or from the identifier (I).

<table>
<thead>
<tr>
<th>Threats in fixed scenario</th>
<th>Additional threats in mobile scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linking of VIDs</strong></td>
<td></td>
</tr>
<tr>
<td>LinkF</td>
<td>LinkM(1)</td>
</tr>
<tr>
<td>identical data in context of VIDs</td>
<td>identical behavior of VIDs observed by identical patterns of identical data or events</td>
</tr>
<tr>
<td><em>Example:</em> identical identifier, identical locator</td>
<td><em>Example:</em> change from identical old locator to identical new locator</td>
</tr>
<tr>
<td>LinkM(1)</td>
<td></td>
</tr>
<tr>
<td>identical behavior of VIDs observed by similar patterns of data or events</td>
<td></td>
</tr>
<tr>
<td><em>Example:</em> simultaneous locator changes with unknown locators</td>
<td></td>
</tr>
<tr>
<td>LinkM(2)</td>
<td></td>
</tr>
<tr>
<td>identical behavior of VIDs observed by similar patterns of data or events</td>
<td></td>
</tr>
<tr>
<td><em>Example:</em> simultaneous locator changes with unknown locators</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Inference of personal information</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>InfFI</td>
</tr>
<tr>
<td>inference from the identifier</td>
</tr>
<tr>
<td><em>Example:</em> home network of the VID if the identifier is a Mobile IP Home Address</td>
</tr>
<tr>
<td>no additional inference from the identifier</td>
</tr>
<tr>
<td>InfFL</td>
</tr>
<tr>
<td>inference from a single locator</td>
</tr>
<tr>
<td><em>Example:</em> location of the user behind a VID at the communication time</td>
</tr>
<tr>
<td>InfML(1)</td>
</tr>
<tr>
<td>inference from several locators</td>
</tr>
<tr>
<td><em>Example:</em> location trace of a user behind a VID over a period of time</td>
</tr>
<tr>
<td>InfML(2)</td>
</tr>
<tr>
<td>inference from user behavior by locator changes</td>
</tr>
<tr>
<td><em>Example:</em> inference of activity by rate of locator changes</td>
</tr>
</tbody>
</table>

Table 2.1: Threats to VIDs by mobile IP communication

Generally, it is possible that an attacker gains a precise knowledge—a VID link or personal information—as well as an imprecise knowledge about a VID. A precise inference is possible, if obvious data is visible to the attacker, e.g., link two VIDs by observing simultaneously the
same locator. Often, this knowledge cannot be inferred precisely but only imprecisely with a certain probability, e.g., the locators of two VIDs are encrypted but simultaneous updates of the locator can be observed. Then, only a suspicion is risen in the attacker. Together with knowledge outside the communication system it is often possible to gain certainty. It is also possible to substantiate a suspicion by observing several states of the mobile system.

3 Threat Analysis of Existing and Proposed Systems

In this chapter, existing systems for mobility management and privacy-aware communication are evaluated when used with multiple VIDs in a mobile environment. The evaluation is done with respect to the threats derived in chapter 2. Aspects not being related to these threats, e.g., performance issues, are not in the scope of this paper.

In literature, there are many systems for anonymous communication proposed. In this paper, it is focussed on typical solutions directly aiming at IP, i.e., GSM/UMTS specific solutions are not considered. Moreover, systems aiming at specific applications only, e.g., WWW, E-mail, file sharing, are not considered. Many evaluated systems originally aim at different objectives than those of this paper, e.g., unobservability of communication. They all fulfill well their primary purpose for which they were invented. In this paper it is analyzed how they could be used to protect against the specific threats in a mobile scenario with multiple VIDs per user, which is a new emerging scenario that was not yet considered in those systems. Moreover, several systems are not specifically proposed for mobile communication. Those systems are combined with Mobile IP for the evaluation.

The addresses used in the systems are evaluated regarding the locator threats as well as regarding the identifier threats. For the evaluation, the addresses are interpreted according to the roles they serve in the concrete scenario. Whenever suitable, the terms locator and identifier are used to emphasize the respective role. Apart from that, the terminology of Mobile IP–which will be introduced below–is used.

In the discussion, protection regarding communication partners and regarding the mobility management system as potential attackers is evaluated. As one main baseline of this paper is the trend towards open communication systems, different entities of the system may be provided by different parties.

For the discussion, the systems are classified into four categories: Mobile IP as starting point and systems providing a similar level of protection, systems providing sender anonymity, hierarchical systems and systems providing some level of protection regarding mobility management entities.

The evaluation is started with Mobile IP which is briefly described in order to introduce the terminology and abbreviations for the following discussions. Mobile IP is an important starting point as its principles are underlying many other proposals and those proposals aiming at fixed communication are combined with Mobile IP for the evaluation.

Class 1: Mobile IP and Systems Providing Similar Protection

Figure 3.1 shows the basic scenario of Mobile IP. A mobile device, called the Mobile Node (MN), has two IP addresses. A fixed one, called Home Address (HoA) assigned from the user’s home network and a variable one, called Care-of Address (CoA) assigned from the respective foreign network the MN is roaming in. The HoA is necessary for reachability by communication partners called Correspondent Nodes (CNs). When the MN is outside its home network, a so-called Home Agent (HA) receives the packets destined to the MN’s HoA (1) and forwards
them in an IP-in-IP tunnel to the MN’s current CoA (2). For this to work, the MN always has to send its new CoA to the HA (3) whenever the CoA changes. Usually, the MN also sends its new CoA to Correspondent Nodes it is currently communicating with (4), so that these can send further packets directly to it omitting the HA on the path. This is called route optimization. Packets originating from the MN and directed to the Correspondent Node can always be sent directly (4). Conceptually, it is also possible to use a (bidirectional) tunnel between the MN and the HA for both directions. This allows to conceal the CoA against Correspondent Nodes. For the evaluation, it is assumed that such a bidirectional tunnel exists.

![Diagram of Mobile IPv6 scenario](image)

**Figure 3.1: Mobile IPv6 scenario**

From the mobility management’s perspective, HoAs principally serve as identifiers and CoAs serve as locators. Due to the nature of IP addressing, the HoA also indicates implicitly the location of the Mobile Node’s fixed presence. Thus, it serves also as locator to the Mobile Node’s home network. In contrast to that, the CoA actually indicates the Mobile Node’s current point of attachment. Thus, the HoA is considered as (non-anonymous) identifier in the privacy evaluation. In order to use Mobile IP with multiple VIDs, a separate identifier as well as a separate locator is used. Multiple locators per device can be realized by using virtual interfaces. This is a loosening of the simplification of section 2. The simplification that the user’s device only has one physical interface still holds. This approach will be followed in the discussion of all systems not supporting VIDs by design.

If a Correspondent Node is in contact with two VIDs of the same user, it sees their identifiers. Thus, protection against threat LinkF depends on the size of the group of potential users of the home network. If it comprises many users, the link between both VIDs will be very imprecise. If only a few users belong to the home network the link will be more precise. The Home Agent can link several VIDs in the fixed case by an observation of similar locators. This is due to the fact that all locators of one device are from the same subnetwork. Again, the precision of the link depends on the size of the subnetwork’s potential user group, i.e., on how many users can potentially have the observed CoAs assigned.

Against the threats LinkM(1) and LinkM(2) Mobile IP protects regarding Correspondent Nodes as they do not see the locators which are the dynamic property. As for each VID a separate locator is used, it is not possible to link several VIDs by observing the identical locator in their context. Thus, protection against LinkM(1) regarding the mobility management is achieved, too. But the Home Agent can observe similar patterns of locator changes in the context of all VIDs of a user. Thus, Mobile IP does not protect against the threat LinkM(2) regarding the mobility management as attacker.

---

1. From the perspective of the foreign network, the CoA serves also as identifier saying which device is addressed.
2. Remember, that a bidirectional tunnel between Mobile Node and Home Agent is assumed.
As both potential attackers—the Correspondent Nodes and the Home Agent—see the identifier, both can infer knowledge from it. Thus, Mobile IP does not protect against threat InfFI.

The locators in contrast are shielded against the Correspondent Nodes but visible to the Home Agent. Thus, Mobile IP protects against threats InfFL, InfML(1) and InfML(2) regarding Correspondent Nodes but not regarding the mobility management.

**Systems Providing Similar Protection as Mobile IP**

The systems evaluated in the following offer similar protection like Mobile IP against the threats of Table 2.1. This is because their original purpose is not protection of multiple VIDs.

The Host Identity Protocol (HIP) [14] is a proposal aiming at the integration of mobility, multi-homing and security in terms of signalling authorization. Its principle is an explicit split of the two concepts of locator and the identifier. The identifier is realized as a cryptographic public key. In order to translate the identifier into a locator, i.e., in an IP address, an address discovery service is used. For mobility support, a so-called Forwarding Agent is foreseen. It receives packets destined to the Mobile Node’s fixed locator and forwards it to the current locator. Mobile Nodes are able to signal their current locator to the Correspondent Nodes. According to the assumption of the bidirectional tunnel used in the evaluation of Mobile IP this locator update at the Correspondent Nodes is not assumed here in order to improve privacy protection.

Regarding the threats of Table 2.1, the only difference to Mobile IP is the use of public keys as identifiers, which contain no sensitive information as such. But the Correspondent Nodes themselves resolve these identifiers in the Mobile Node’s fixed IP address and thus, also see this address. This IP address is conceptually similar to the HoA serving as identifier in Mobile IP. It points to the Forwarding Agent which is assumed to be in the user’s home network. Thus, the evaluation result is the same as with Mobile IP.

The Non-Disclosure Method [15] is an extension to Mobile IP that shields the locator updates between the Mobile Node and its Home Agent against observation by third parties. This is achieved by so-called Security Agents working similar to cryptographic Mixes [16]. Protection regarding the Home Agent and the communication partners as potential attackers is not considered. Thus, with respect to the threats of Table 2.1, the evaluation results are the same as with plain Mobile IP.

The Freiburg Location Addressing Scheme (FLASCHE) [17] aims at protection against the link of several actions of a Mobile Node and of those actions link to the user’s device. This is achieved by not using the same identifier over a considerable period of time. Thus, temporary identifiers are used. They consist of a random part as well as a part containing the Mobile Node’s current location. FLASCHE relies on the frequent change of the location, because thus, the identifier changes frequently.

This approach is only designed for communication initiated by the Mobile Node itself. Backward traffic can only be received as long as there are only minor location changes. For reachability, the system can be combined with Mobile IP, i.e., using FLASCHE between the Home Agent and the Correspondent Nodes. This in turn undermines the protection, as the identifier does no longer change. Thus, in the general scenario of communication being initiated by the Mobile Node and Correspondent Nodes, the same protection is achieved as with Mobile IP.
Class 2: Systems Providing Sender Anonymity

In the following, systems are evaluated that originally aim at unobservable communication in a fixed scenario. Some of them are part of a comprehensive solution also providing, e.g., application data filters. All of them provide for sender anonymity, which is the only property being relevant in the evaluation of this paper.

The conceptual idea behind Onion Routing [4] is to use cryptographic Mixes. Thereby, no observer can see which nodes are communicating. Moreover, the recipient does not see the sender’s address. Backward traffic is basically supported, but only from nodes that have been contacted previously and that have been given an anonymous identifier.

Crowds [6] aims at sender anonymity and unobservability regarding local eavesdroppers and the entities of Crowds. Several entities called Jondos are in the communication path between the sender and the recipient. Each Jondo forwards the packets with a certain probability to the final recipient and to another Jondo otherwise. Thus, neither any Jondo nor the final destination knows, which Jondo originated the packet¹. Again, it is possible to receive information anonymously from previously contacted nodes.

Hordes [5] relies on the same principle for traffic sent by the Mobile Node. For backward traffic, multicast addressing is used, thus hiding the originator in the multicast group.

In order to provide reachability and mobility support, this section’s systems can be combined with Mobile IP. For that, the systems are applied between the Home Agent and the Correspondent Nodes. For Onion Routing and Crowds, it must be distinguished between Correspondent Nodes, that have been contacted previously and between Correspondent Nodes that initiate a first contact. Latter ones must know a permanent identifier of the VID which leads to the Home Agent, i.e., a HoA. Previous ones can send packets in reply to an anonymous contact, thus knowing only an identifier not containing any sensitive information.

Regarding Correspondent Nodes not being contacted previously and regarding the entities of the mobility management the evaluation results are the same as for Mobile IP. The only difference is for Correspondent Nodes that have been anonymously contacted by the Mobile Node before. Regarding them, the systems additionally protect against threats LinkF and InfFI as these Correspondent Nodes do not see the HoA as identifier containing sensitive information which can lead to a link of VIDs or inference of personal information. In Hordes this applies for all Correspondent Nodes as the HoA can be hidden in the multicast group. It is not foreseen by the authors that this fixed address can be accessed by Correspondent Nodes without a prior contact by the Mobile Node but conceptually, it is possible.

Class 3: Hierarchical Systems

In this section, systems are evaluated, that consist in principle of a hierarchy of entities which are similar to a Mobile IP Home Agent. The lower the entities are placed in the hierarchy, the smaller is the network they are responsible for. The networks are smaller in terms of geographical extent as well as in terms of users potentially being attached to them.

Figure 3.2 shows the scenario of Hierarchical Mobile IP [17]. The highest level entity (HA) is similar to a Home Agent. Additionally to that, there may exist several levels of so called Mobility Anchor Points (MAPx). These agents build up the mobility management system. A Mobility Anchor Point serves as a kind of a local Home Agent for the Mobile Node’s current roaming area. The Mobile Node does not only have a HoA in the Home Agent’s network and a

¹. The sender as part of the Crowds system is also a Jondo.
CoA in the visited foreign network but also a so-called Regional CoA (RCoA) in the Mobility Anchor Point’s network. The currently assigned address is called LCoA, On-link CoA, and resembles a common CoA. Therefore, the term CoA is used in order to avoid confusion. On local movement, only the binding between the changing CoA and the RCoA has to be updated at the Mobility Anchor Point (1). Only if the area of a Mobility Anchor Point is left, the binding between the HoA and the new RCoA at the Home Agent must be updated (2) additionally to the update of the CoA (3). Thus, local movements are transparent to the Home Agent which reduces signalling overhead. This was the original intention of Hierarchical Mobile IP.

![Figure 3.2: Hierarchical Systems](image)

The Home Agent is fixed in the user’s home network like in standard Mobile IP. Hence, the HoA contains sensitive information about the user’s home, because of its implicit functionality as a locator of the Mobile Node’s fixed presence. The current Mobility Anchor Points at the different levels are determined due to the user’s current location and change when the user moves. Thus, RCoAs and the CoA contain sensitive information about the user’s current location.

Like the HoA, the addresses used between the agents, i.e., the RCoAs, serve as both, identifier and locator. For the sending entity of the higher level they are a locator, i.e., a pointer to the location where to send the packet. For the receiving entity on the lower level, they serve as an identifier, i.e., an indicator which Mobile Node is to be contacted.

The relevant property of the system for this paper is the potential for privacy protection. Only the lowest level Mobility Anchor Point knows the Mobile Node’s current exact locator. Entities at higher levels only see addresses serving as intermediate locators, i.e., RCoAs, to the next lower Mobility Anchor Point’s network. Thus, these intermediate locators contain less accurate information than the Mobile Node’s current exact locator, i.e., the CoA.

The system described in [18] is conceptually similar to Hierarchical Mobile IP. Here, the hierarchically organized entities are called Mist Routers. They build up a tree-like overlay network. Between the Mist Routers, link identifiers and pseudonymous handles are used for path finding. No entity knows the path through all Mist routers used by a VID. Only the top level Mist Router knows the VID’s identifier, while only the lowest level Mist Router knows the current exact locator. Higher level Mist Routers can derive sensitive information from knowledge of the next lower level Mist Router which is indicated by the outgoing network link. This is similar to an intermediate locator like the RCoAs used between the agents in Hierarchical Mobile IP.
In the hierarchical systems evaluated in this section, protection regarding Correspondent Nodes as potential attackers is the same as with Mobile IP. This is, because the path from the Correspondent Nodes to the Home Agent or top level Mist Router is identical. Correspondent Nodes can see the identifier but not the current locator.

For evaluation of the threats regarding the mobility management system as potential attacker—which here consists of several entities—it is assumed that different hierarchical entities are operated by different parties. This is necessary to distribute the sensitive information among several parties. It is a reasonable assumption considering the multiple provider situation in future communication systems and the fact that the top level entity resides in the user’s home network whereas the lower level entities reside near the user’s current location. Protection against threat LinkF depends on the size of the intersection set of the potential groups of users of the respective entity and of the next lower level entity. This is because all VIDs of a user are using the respective entity as well as the same next lower entity. The more users belong to this intersection set, the larger is the anonymity set in which the VIDs of the considered user are hidden.

As again for each VID a separate locator is used, it is not possible to link several VIDs by observing the identical locator in their context. Thus, protection against LinkM(1) regarding the mobility management is achieved. Against threat LinkM(2) this protection is not achieved. The only difference to plain Mobile IP is that the different entities observe different kinds of behavior patterns. The lower the level, the more often a change can be observed. Changes occur only in the lowest level entity that is not changed due to the user’s movement. The changes are transparent to higher level entities whereas lower level entities are changed themselves.

Considering protection against threat InfFI regarding the mobility management system as attacker, the highest entity in the hierarchy sees the identifier. Moreover, the entities of each level see the respective intermediate identifier used by the entity above in the hierarchy which does not contain any sensitive information the lower level entities not already know. This intermediate identifier serves as an intermediate locator for the upper entity. Considering threats InfFL, InfML(1) and InfML(2), each hierarchical entity knows the locator pointing to the next lower entity. Moreover, the lowest entity knows the Mobile Node’s current locator.

Class 4: Systems Protecting Against Mobility Management Entities

The systems evaluated in this section aim at concealing the link between the HoA and the CoA regarding the mobility management system as well as the link between Home Agent and the Mobile Node regarding external observers by use of Mix-like entities. They assume different parties running the different entities. The Home Agent knows the VID’s identifier (the HoA) but not its locator (the CoA). The last entity of the chain in turn knows the locator but not the identifier. In principle this is a split of the Home Agent’s two functions as fixed presence of the Mobile Node for reachability and as reference to the current location of the Mobile Node.

Mixed Mobile IP [19] proposes an approach in which there are two Mix-like entities between the Home Agent and the Mobile Node. Figure 3.3 shows the scenario. The Correspondent Node (\textit{CN}) sends packets to the \textit{Home Agent}. The \textit{Home Agent} does not know the current loca-
tor but forwards the packets to Mix 1. Mix 1 in turn forwards the packet to Mix 2 which knows the locator and finally delivers the packet. Originally the approach is designed for Mobile IPv4 with a so-called Foreign Agent in the currently visited network being the endpoint of the IP-in-IP tunnel from the Home Agent. This Foreign Agent knows the user’s locator. In principle, the approach can be transferred to Mobile IPv6 by the last Mix knowing the user’s locator. This is the scenario evaluated here.

Flying Freedom [20] extends the Freedom system [21] towards handling mobility. The Freedom system is an overlay network using Mix-like entities, so-called Anonymous Internet Proxies (AIPs), for pseudonymous communication. Between the AIP serving as Home Agent and the last AIP knowing the locator there may be a chain of AIPs of arbitrary length. For the discussion it is assumed that the AIP serving as Home Agent is located in the user’s home network. The system supports multiple VIDs per user and shields the current locator against communication partners. Each VID has an IP address being part of the network in which the AIP acting as Home Agent is located. This is in accordance with the approach taken in the other systems in which each VID has a HoA.

There exist two scenarios how the systems of this section can be used with VIDs. In the first one, there is one Home Agent (or AIP serving as such) and one Mix-like entity knowing the locators for all VIDs. From the point of view of the relevant threats in this paper this scenario is similar to the hierarchical approaches but the Mix-like entities can be located anywhere, thus not containing sensitive location information. In the second scenario, different Mix-like entities are used for different VIDs. The latter scenario requires a lot of resources in terms of Mix-like entities and overhead of signalling, as a locator change has to be signalled to each entity being responsible for one VID’s locator. Especially with regard to mobile communication across wireless links the latter overhead could be a restriction. Moreover, several entities see the locators and can derive sensitive knowledge from them. In the following, both scenarios are evaluated against the threats of Table 2.1.

The results regarding protection against Correspondent Nodes are the same in both scenarios. Protection against threat LinkF is the same as with Mobile IP. Against LinkM(1) and LinkM(2) the system protects as the Correspondent Nodes do not see dynamic locators. While Correspondent Nodes can derive sensitive knowledge from the identifier (threat InfFI) they cannot derive sensitive knowledge from the locators (threats InfFL, InfML(1), InfML(2)).

Regarding the mobility management system, in the first scenario–with the same entities for all VIDs–the first entity in the row sees the identifier. Thus, no protection against InfFI is achieved with respect to this entity. Moreover, protection against LinkF depends on the size of the intersection of the potential user groups of the respective entity and of the following entity in the chain or of the foreign network respectively. As for each VID a separate locator is used, it is not possible to link several VIDs by observing the identical locator in their context. Thus, these systems protect against LinkM(1) regarding the mobility management. As the last entity of the chain sees the locators and their changes, no protection against LinkM(2), InfFL, InfML(1) and InfML(2) is achieved regarding this entity.

In the second scenario–with different entities for different VIDs–protection against LinkF regarding the Home Agent depends only on the user group it is serving. As no Mix-like entity serves more than one VID, protection against LinkF is achieved regarding them. For the same reason, protection against LinkM(1) and LinkM(2) is achieved. Protection against LinkM(1) and LinkM(2) is also achieved regarding the Home Agent, as it does not observe any dynamic locator.
The Home Agent can infer knowledge from the identifier, while the respective last Mix-like entities in the chains can infer knowledge from the locators. Thus, no protection against InfFI regarding the Home Agent and no protection against InfFL, InfML(1) and InfML(2) regarding the last Mix-like entities in the chains is achieved. Note, that in this scenario there are several last entities in the chain regarding which the system is vulnerable against these threats.

Table 3.1 summarizes the evaluation results. In each row there are the results regarding one threat. The columns contain the results regarding the different groups of systems. For each system the protection regarding Correspondent Nodes and regarding entities of the system is listed. In the last column, there are two results—one for the first scenario with the same Mix-like entities for all VIDs of one user and one for the second scenario with different entities for different VIDs of one user. A "+" means that protection is achieved while a "−" means a vulnerability, at least against some entities of the mobility management system.

Like can be seen in Table 3.1, none of the existing communication systems can protect VIDs against all of the threats derived in chapter 2. This is mainly due to the reason that most of them do not aim at multiple VIDs. Moreover, the majority does not aim at a mobile scenario or the requirement of reachability of a pseudonymous user. Therefore, most of the systems do not provide sufficient protection against the linking threats, especially in the mobile scenario.

<table>
<thead>
<tr>
<th>Threat</th>
<th>Mobile IP and similar systems</th>
<th>Systems providing sender anonymity</th>
<th>Hierarchical systems</th>
<th>Systems protecting against Home Agent (scen. 1 / scen. 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LinkF</td>
<td>_ a</td>
<td>_ b</td>
<td>_ a/b/c</td>
<td>_ b</td>
</tr>
<tr>
<td>LinkM(1)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>LinkM(2)</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>InfFI</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>InfFL</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>InfML(1)</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>InfML(2)</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
</tbody>
</table>

Table 3.1: Results of the evaluation

a. Depends on size of potential user group of home network
b. Depends on size of potential user group of foreign network
c. Protection regarding previously contacted Correspondent Nodes can be achieved. In Hordes this is possible for all Correspondent Nodes.
d. Depends on size of intersection of potential user groups of respective entity and of next level entity
e. Depends on size of intersection of potential user groups of respective entity and of next entity in chain
f. Last entity in chain sees locators and their changes but not identifier
g. First entity sees identifier but not locator
4 New Approach for Privacy-Preserving Mobile Communication with Multiple VIDs

In this chapter an outlook to a new conceptual approach as presented in [22] is given. In Figure 4.1 the basic scenario is sketched.

A common vulnerability of the systems evaluated in chapter 3 is inference of sensitive information from the identifier (InfFI). In this new approach, the identifier, which resembles a Mobile IP Home Address is not from the user’s home network, but from an arbitrary subnetwork of the Internet, thus not containing any sensitive information about the user. In order to protect against LinkF regarding Correspondent Nodes, the identifiers of different VIDs are chosen from different subnetworks each having a Home Agent (HA1-x) running. Thus, the user’s fixed presence is no longer related to the real, physical home and the user has several different fixed presences.

Again, it is assumed, that different subnetworks are operated by independent parties to distribute the sensitive information among them. As outlined in the introduction, a large number of providers is realistic in a system with an open philosophy.

Several proposed systems are prone to linking attacks by the mobility management system’s entities. In order to eliminate these threats in the proposed approach, separate contexts are retained for each VID throughout the packet’s path to the Mobile Node. The packets of each VID are forwarded to a different Home Agent of the second level (HA2-x). In order to protect against inference of sensitive information by the HA2-x, these agents are also arbitrarily distributed in the network. Thus, they don’t contain any information about the user’s current location in contrast to the hierarchical systems of class 3. Moreover, by having two entities in a row, it is achieved that no entity knows both, the identifier and the locator.

As in the systems of class 4 when used with the second scenario described above, the separation of VID contexts duplicates the sensitive knowledge of the locator. Thus, protection of the locator becomes even more important than in simpler systems with only one entity knowing it.

As again different CoAs are used for different VIDs, protection against LinkM(1) considering the mobility management as potential attacker is assured. Moreover, the locator is invisibly stored among several parties. Only if it is needed, i.e., in case of a packet arriving for the respective VID, the locator is recombined. Technically, this can be achieved by secret sharing techniques [23], in which the sensitive information is split into several pieces which are all necessary to recombine the information but which do not contain any sensitive information

1. The user can trade-off this gain on privacy for a gain on performance when choosing the agents near to his current location.
2. In order not to disclose the current CoA when updating a share, the user must use an anonymous tunnel here. As this is a client-initiated communication there exist many solutions which can be used.
when being observed alone. The user can have the choice in how many pieces the secret is spread thus controlling his level of protection. Thus, protection against LinkM(2), InfFL, InfML(1) and InfML(2) can be perfectly achieved during silent phases of a VID which often is the majority of the time.

To improve protection against the dynamic threats LinkM(2), InfML(1) and InfML(2) during communication times, the entity being able to recombine and thus, observe the locator \((HA2-x)\) is changed frequently. Thereby, it is firstly not possible to gain a long trace of locators from which sensitive information can be inferred and secondly it is not possible to substantiate an imprecise suspicion about two VIDs assumed to be linked.

As compared to the existing systems of class 4, the number of necessary entities can be decreased to a more acceptable number. This is possible by the invisible storage of the sensitive locator which allows an entity of the mobility management system to serve several VIDs of a user simultaneously as long as only one of them is communicating. Even if they communicate simultaneously, they can be linked only imprecisely—as different CoAs are used for different VIDs—and this link cannot be substantiated by several observations as the VIDs change their serving entities. The probability of this case to happen depends on the concrete usage scenario in the same way as the preciseness of the observed link depends on the anonymity set of the concrete scenario, i.e., how many different users can potentially have the observed similar CoAs. This allows the user to choose a trade-off between the number of necessary entities—which have to be paid in a real scenario—and the level of protection.

In the following a summary of the privacy evaluation of the sketched approach is given. Conceptually, it starts from the achievements of the systems of class 4 when using them in the second scenario as discussed. Like there, no entity is knowing both, the identifier and the locator. The results regarding the concrete threats are as follows:

- **LinkF**: Different HoAs from different subnetworks are used for different VIDs. Thus, these identifiers are completely unrelated. This protects regarding both, the Correspondent Nodes and mobility management system as potential attackers.

- **LinkM(1)**: Different CoAs are used for different VIDs. Thus, no entity can observe identical locators of several VIDs.

- **LinkM(2)**: If it can always be assured that one HA2-x only serves one VID of a user at a time, similar behavior cannot be observed in the context of several VIDs. In case of one HA2-x serving several VIDs of a user, it can happen that an imprecise link can be drawn by observation of similar CoAs if the VIDs communicate simultaneously. Because of the regular change of HA2-x entities, this imprecise link cannot be substantiated by a long observation. An imprecise link can also be drawn in case of simultaneous updates of the invisibly stored CoA, i.e., the shares of the secret sharing scheme. As the only property rising suspicion, is the identical time of these updates, those links are very imprecise. Again, a substantiation of the suspicion can be avoided by a regular change of the entities.

- **InfFI**: The identifier is anonymized by being from a subnetwork not related to the user. This protects against both potential attackers, the Correspondent Nodes as well as the mobility management system.

- **InfFL**: Inference of sensitive information about a VID from a locator is only possible if the respective VID is currently communicating, i.e., its CoA is recombined by a HA2-x. This threat cannot be fully prevented as the locator is unconditionally necessary to deliver packets. But the threat is minimized the best possible by storing the locator invisibly as long as
no communication takes place. Moreover, the HA1-x cannot infer any knowledge from an intermediate locator pointing to a HA2-x as this agent can be located anywhere. As the Correspondent Nodes do not see the locator, it is also protected against them.

- InfML(1): This threat is diminished by changing the HA2-x. Thus, no HA2-x has a long trace about a VID from which sensitive information can be inferred. Again, the HA1-x cannot infer any knowledge from the intermediate locators pointing to the HA2-x agents as those can be located anywhere. Moreover, they can be changed arbitrarily, thus containing no information about the unique behavior of the user. As the Correspondent Nodes do not see the dynamic locator, it is also protected against them.

- InfML(2): see InfML(1)

As this approach is based on distribution of sensitive knowledge to different entities, it principally has drawbacks in terms of availability that can partly be alleviated by redundant schemes in which only m of n pieces are necessary to fulfil the functionality. Moreover, it causes performance costs in terms of packet delay and signalling overhead.

As different users have different preferences regarding privacy or performance, it is important that the approach gives the user the control on this trade-off. The proposal allows a configuration like a class 3 system, even with the same CoA for all VIDs, thus gaining the best performance–due to the hierarchical nature even better than plain Mobile IP–or it allows a configuration like described in this section in order to achieve maximum privacy. The quantification of this trade-off is out of the scope of this paper which focusses on the concepts from a privacy point of view. The same applies for authentication, authorization, accounting and charging schemes which can principally be solved, e.g., by anonymous credentials like proposed in[24].

5 Conclusion and Future Work

In this paper, the threats to VIDs caused by mobile IP-based communication were thoroughly analyzed. It turned out that these are basically the threat of links between several VIDs of one user and inference of personal information of the user behind a VID. It was pointed out, that dynamics of mobile communication adds largely to the complexity of the challenge of solving future privacy needs.

Afterwards, existing proposals for anonymous communication were evaluated regarding the derived threats. Therefore, they were classified into four groups. Mobile IP and systems providing the same protection, systems providing sender anonymity, hierarchical systems, and systems achieving some protection against the mobility management itself. The level of protection achieved rises with each group, but none of the evaluated systems protects against all threats.

With the weak points of the existing systems in mind, an outlook to a new approach protecting VIDs in mobile communication was given at the end. It allows the user a large freedom in controlling the trade-off between performance, scalability and privacy. In the future, this approach will be extended and detailed. As privacy protection always is a trade-off with performance, the exact quantification of this will also be a future working area. Therefore, a performance analysis has started.
References


