

Privacy, Control and Internet Mobility

Tuomas Aura and Alf Zugenmaier¹

Microsoft Research, Cambridge, United Kingdom
tuomaura@microsoft.com, zugenmaier@docomolabs-euro.com

Abstract. This position paper explores privacy issues created by mobile and wireless Internet access. We consider the information about the users identity, location, and the services accessed that is necessarily or unnecessarily revealed observers, including the access network, intermediaries within the Internet, and the peer endpoints. In particular, we are interested in data that can be collected from packet headers and signaling messages and exploited to control the users access to communications resources and online services. We also suggest some solutions to reduce the amount of information that is leaked.

1 Introduction

As laptops and hand-held devices replace stationary desktop computers as the mainstream personal computing device, the average Internet user and host become increasingly mobile. The mobile devices stay connected to the Internet via business and domestic broadband networks, wireless hotspots, and cellular data links, bringing closer the dream of universal connectivity.

The connectivity is implemented with the help of a number of autoconfiguration and mobility protocols. The user, however, is not exposed to all this complexity. The devices do their best to hide the protocol details from the user and to provide seamless universal access to online services. We can expect them to get better at this as the mobile access technologies mature. Ideally, the user remains connected to the same services regardless of the location and access network.

While transparent connectivity is a powerful tool, it is worth considering the potential adverse consequences that it has for the users. One such consequence is that the users expectations of privacy may not match what goes on at the network protocol level. As the users and their mobile devices move seamlessly from one access scenario to another, packets traveling over the access links and across the Internet may reveal information that the user would prefer to keep confidential. The same information may also be used to control the way the user communicates and to differentiate the services and prices offered to individual users.

The information that we consider in this paper is limited to the users identity, location and the services they access. Protection of critical application data is

¹ Alf has moved to DoCoMo Euro-Labs, Munich, Germany.

outside the scope of our discussion. That is, we are interested only in information that is carried in the packet headers or revealed by standard signaling messages. IPv6 is particularly important for our discussion because many new mobility and security solutions have been designed for the next-generation Internet protocol. Nevertheless, we also consider IPv4 where appropriate because IPv4 will be around for quite a while and mobile hosts will have to alternate between the two technologies to gain maximal access to the Internet. It is also interesting to compare the new and old protocols from the privacy point of view. It should be noted that we limit the discussion to the most common standard Internet protocols and medium access protocols. It is quite possible that proprietary protocols and extensions either reveal additional information or help to hide it.

The privacy of mobile users has previously received quite a lot of attention. See, e.g., [4] for identity protection and [5,6,10] for location privacy. We refer the reader to [3] and [17] for comprehensive overviews. This paper adds the control aspect to the discussion. We provide novel viewpoints to the privacy implications of various networking technologies. We consider practical privacy enhancements to Internet protocols rather than theoretically strong anonymity mechanism. See, e.g., [15] for an overview of strong anonymity.

The rest of the paper is organized as follows. Sections 2-3 discuss the kinds of information that may be collected, the entities that may collect it, and the reasons why they might want to do so. In Section 4, we explain some common mechanisms for privacy protection and discuss their limitations and potential improvements. Section 5 concludes the paper.

2 Information that may be gathered

When thinking about their privacy, users are typically worried about two kinds of exposure. First, users do not want to reveal names and other identifiers that enable someone to uniquely identify them. Second, users are wary about exposing personal information, such as their address, sex, age, job, and salary. Users are particularly reluctant to reveal the personal information if it can be linked to the unique person. These concerns arise from experience: most personal information is currently collected for marketing purposes and the goal of the collectors is typically to link a person to an address and other demographic information.

In mobile Internet access, it is also possible to identify the user, either by name or by another permanent identifier. If the user is paying for network access or online services, the service provider typically knows the users name and other personal details. Otherwise, Internet protocols reveal many identifiers that can be used to uniquely identify the device or user, although rarely the users name or other personal details. Usually it is sufficient to find out the identity of the device used because most mobile devices belong exclusively to one user, or to a small group of users. The available identifiers include:

- device identifier, e.g., the network interface MAC address,
- host identifiers used by mobility protocols, e.g., the Mobile IP home address or the HIP host identifier, and

- application-layer identifiers, e.g., HTTP cookies or a media player serial number.

Device identifiers are only visible at the link layer while host identifiers are typically seen both at the local link and by the mobiles correspondents across the Internet. The application-layer identifiers are intended only for the correspondents. They may be visible at the local link if the packets are not encrypted, although sniffing is practical only for standardized protocols and identifiers.

The information gathered about users is likely to be somewhat different from the usual demographic data in which advertisers are interested. Instead of an address or a salary range, the network elements can learn the information in network signaling messages and in packet headers. The information available to observers from such sources includes:

- the mobiles location, i.e., current IP address,
- protocols and applications used,
- web sites visited and other online services contacted, and
- VPN gateways of the users organization.

In this paper, we consider the users privacy to be at risk when an observer is able to connect a communication event to a particular device or user. This information may, in fact, be more interesting to the network operators or advertisers than demographic data about the user would be. They can, for example, tell apart business and entertainment use. Moreover, even if the observer cannot link the particular identifier or event to a person, it may be sufficient to track pseudonymous users and their communication habits.

3 Common observers and their motives

Traditionally-recognized threats to privacy come from such observers as direct marketers, nosy individuals, and overbroad law enforcement access. The common feature of these observers is that their main goal is to gather information about the target individual or group. These threats are not fundamentally different from the problems that existed in the pre-Internet society. The threats to privacy have been highlighted in the Internet age because digital data formats and computer networks enable more efficient data aggregation and because the users habits have not yet adjusted to the new medium.

Discussions on Internet privacy often focus on one or more traditional data collectors and on potential offline misuse of the information collected online. We, however, argue that this focus is preventing us from seeing another serious threat to user privacy and freedom: the use of the online information for controlling the users access to communications resources and online services. This threat arises from the pressure for network and service operators to maximize their revenue by price and service differentiation and market segmentation. Clearly, the operators can charge more if they know how much each individual connection or data packet is worth to the user. In order to be able to do so, the operators

need detailed information about the protocol and services that each mobile node is using.

Fixed-line Internet access is usually priced by the bandwidth or by the time spent online regardless of which services the users access. The same is not true for GPRS and other mobile networks. There are complex pricing structures in place where the cost of Internet access depends on the protocol, application, and server accessed. Telephone-network operators sometimes block access to online services that compete with their own offerings. Mobile phone operators routinely charge different prices for IP traffic depending on the application-layer protocol, the server contacted, and the location of the mobile. The effect of such restrictions is twofold. First, the operators can directly control the users Internet access without the need to first gather demographic information and then using that information for market segmentation. This enables much more accurate price and service differentiation, to the extent that each new protocol and destination address is allowed only if the user pays a separate subscription for it. Second, the users find themselves unable to access arbitrary services on the Internet, including new and innovative ones, until the network operator has a service plan and a pricing structure in place for them. Thus, mobile users are unable to reap the full benefits of universal Internet access.

To summarize, we are concerned about two kinds of information gathering:

- collection of business, personal or law-enforcement
- intelligence about the users for offline use, and
- using communications metadata to exert fine-grained control over the users online activities.

The network elements where information can be collected include:

- routers and servers in the access network,
- other nodes in the access network,
- name servers and other network infrastructure,
- middle boxes such as NATs, proxies, firewalls, and other traffic filters and shapers,
- mobility infrastructure, and
- peer endpoints, such as web and email servers.

The item that is notably missing from this list is the core network routers. Theoretical and extremely paranoid discussions of Internet privacy sometimes assume that the observer may be anywhere in the network. It is, however, impractical to implement a Dolev-Yao-type omnipresent attacker on the Internet. Instead, the noteworthy observers are either located at key nodes on which the communication depends, or they are at edge nodes such as servers and other mobiles. The former will appeal to casual and covert observers while the latter are the best places for service providers and authorities to establish maximal control.

The mobile user typically accesses Internet via a wireless LAN, wired LAN, or cellular phone network. In a broadcast access network, such as a wireless LAN,

other hosts can listen to the traffic between the access point and the mobile host. Even if the end-to-end communication is encrypted, the packet headers and signaling messages are typically visible to the local nodes.

The nodes in the local network can identify the mobile by link-layer and IP-layer identifiers that are in the unencrypted part of the packet. These include the link-layer address and Mobile-IPv6 home address. These addresses are visible to all nodes on the network if the LAN is unprotected by link-layer encryption, or if the encryption is based on a shared key. Regardless of link-layer encryption, the wireless access router always sees the IP-packet headers and signaling packets that have not been encrypted in the network layer. The router can, for example, learn the IP addresses of the mobiles correspondents. Moreover, the home address (in a Mobile-IPv6 home address option and routing header) cannot be encrypted at the network layer because IPsec security associations between the mobile and its correspondents are identified by the home address.

Wireless LANs that encrypt data on the link-layer using a different key between each mobile and the base station are equivalent to switched wire LANs: only the access router is able to see the IP packet headers. The mobiles link-layer address is still visible to everyone.

Most mobiles access a DNS server provided by the access network, which is typically configured with the DHCP protocol. The DNS server is able to record the names of the online servers contacted by the mobile. Even if the mobile connects to a VPN gateway and uses DNS services via a VPN tunnel, it may still rely on the local DNS server to resolve the VPN gateway name. This means that the DNS server in the access network learns the name of the organization to which the user belongs, and may enable it to identify the mobile with some accuracy. Furthermore, DNS requests are made recursively, which leaks the mobiles approximate location to the remote DNS servers. Thus, even if the actual data connections are forwarded via anonymizing proxies, the source of a DNS request may reveal the mobiles location to the peer endpoint.

In addition to the access routers and name servers, there are other bottlenecks on the network that can see and control all traffic to and from a mobile host. These include NATs, Web proxies, and firewalls and other traffic filters. In many cases, the existence of such middle boxes has the effect of preventing the hosts in the access network from using some features of the Internet protocol suite. For example, a NAT prevents the hosts from acting as servers, except for specifically approved protocols, and a firewall can be used to filter out protocols and destinations not endorsed by the network operator. Together with information acquired from authentication and sniffed signaling, such middle boxes enable fine-grained control of the services afforded to each mobile.

Mobility agents, such as Mobile IP [13] home agents, HIP [12] forwarding agents and SIP [14] registrars, are another point at which location information can be collected. The mobility agents may also be able to differentiate between different applications and, thus, control the mobiles Internet access. Indeed, such control is one of the main functions of a SIP server. SIP servers enable two Internet nodes to find each other even if neither one of them has a permanent IP

addressed. SIP servers are often provided by a network or VoIP service operator. They typically perform accounting functions and, therefore, identify and authenticate the user. Unlike SIP, Mobile IP does not specify a mechanism for differentiating between applications but, in the reverse-tunneling mode, the home agent could implement similar control. Moreover, location information collected by the mobile can be used to vary prices depending on the mobile's location.

The peer endpoints can, of course, also collect information on the mobile node. Since this is the privacy treatment model that has received most attention in the past, it suffices to mention that mobility does not change the threats all that much. Only the location of the mobile is added to the data that the mobile's correspondents can learn about it. Advertising-based services, including most online publications, need to know their customer in order to target advertising, and the location is one valuable additional piece of information. Mapping the IP address to a geographical location is, however, relatively unreliable.

4 Some simple and effective privacy solutions

In this section, we consider simple and inexpensive privacy solutions that exist in current networks. We are interested only in mechanisms that can be deployed locally with little cost in either infrastructure or communications overhead. We suggest ways of using these mechanisms effectively to improve the privacy of mobile users.

4.1 Roaming vs. mobility

Roaming is a term more commonly used in cellular phone networks than in TCP/IP. Nevertheless, it is useful in making a distinction between different types of mobile access. Roaming means that a mobile host gains access to the Internet via some other access network than its usual network. For example, if one takes a laptop computer from work to home and connects to the Internet via the network at one's home, the laptop is roaming there. The reader might wonder what is so unusual about roaming. The answer is that, in mobile phone networks, roaming cannot be implemented without a complex inter-operator charging system and contracts drawn by lawyers.

For the more orthodox type of engineer, roaming is not true mobility. Mobility means that the mobile is moving from one network to another. Moreover, mobility is associated with two technical challenges that the engineer expects a mobility system to solve. First, connections between the mobile should survive the movement, even across the boundaries of physical networks. Second, the mobile should be reachable wherever it goes. That is, other network nodes should be able to contact the mobile. Solving these two problems requires fairly sophisticated technology, such as the cellular mobility management protocols or Mobile IP.

From the privacy perspective, roaming, regardless of its technological inferiority, is more attractive than mobility. Consider a mobile host that attaches

to the Internet via whatever access network is currently available, obtains an address at the access network, and behaves in every way like a local node at that network. It can access many Internet services, such as web servers, without identifying itself or telling anyone that it is a mobile host.

On the other hand, Mobile IP(v6) and other mobility protocols require the user to identify itself to the servers and peer hosts by some permanent identifier, so that the same mobile can be recognized when it reconnects from a new location. Many mobility protocols also require the mobile to register its location in some central or distributed database in order to implement the reachability.

It is clear that user privacy would be better protected by using the mobility protocols only when their features are needed and simple roaming access at other times. The reason why this is rarely done is that it is difficult for the computer to know what level of service the user needs, and how soon the user intends to move. It is also difficult for the IP, transport or middleware layers, where mobility is often implemented, to know whether a particular application requires mobility or just roaming. Quite clearly, it would help to make the applications more aware of mobility, and the protocol stack more aware of the application requirements. It may, however, take a while before the protocol stack layers and applications communicate with each other about privacy requirements.

4.2 Anonymizing proxies and Mobile IP tunneling

Forwarding packets via a proxy hides the individual sender within the group (anonymity set) [2]. The most common type of proxy is an HTTP proxy. They are commonly used at the boundary of corporate networks and hide the sources of individual web requests. This is useful even in a stationary network: an observer on the Internet (usually the web server) knows that a request came from the specific organization but it cannot tell which member of the organization made it.

Some web proxies (e.g., Anonymizer) exist for the sole purpose of anonymizing web access. In that case, they may be located far from the user who sends requests to the anonymizing proxy via an encrypted tunnel. They also try to attract large and diverse groups of users, with the aim of hiding not only the identity of the individual user that makes a request but also the organization from which the request originated.

The way we usually think about the effect of the proxies is that they hide an individual in a group. If the group is large enough, it is impossible to distinguish individual access patterns in the pool of web requests. In a mobile access network, however, the proxies have another effect that is perhaps more significant: a proxy at the access network hides the existence mobile nodes that are connected to the local network behind the proxy.

One can also see a NAT [16] as a proxy that handles individual IP packets. It hides the individual nodes within the network and presents only a single IP address to the Internet. In addition to having the same privacy benefits as a web proxy, a NAT has the effect of hiding the number of IP hosts in the network, as well as the structure of the network behind the NAT. Obviously, the existence

of any mobiles is hidden as well. A limitation for both the NAT and the web proxy is that if the number of nodes behind the proxy is small, the anonymizing effects become weaker. (We will have more to say about NATs and IPv6 below.)

The Mobile IP home agent is also a kind of proxy between the mobile node and the Internet. If the mobile uses reverse tunneling, all packets are routed via that home agent. The effect of this proxy is, however, different from the proxy in the access network. It hides the mobiles location, and even the fact that the mobile is a mobile, while revealing its the home address (i.e., permanent identifier) to the peer endpoints. This may sometimes be exactly what the mobile user wants.

Mobile IP reverse tunneling implements full mobility, as defined in the previous section. So how is it able to hide the mobiles location? The answer is that reverse tunneling is a trade-off between performance and privacy. By sending all communication via the home agent, the path taken by the packets may be much longer than it would be if the packets were sent directly.

To be honest, the original trade-off was between performance and technical complexity but, at least in Mobile IPv6, that is no longer an issue as implementations of route optimization exist. It would be desirable to make the choice between reverse tunneling and route optimization (i.e., direct communication between the mobile and the peer endpoint) depending on the application requirements. Unfortunately, the state of art in Mobile IP(v6) implementations does not yet allow this.

4.3 Mobile IPv6 route optimization and IPsec

IPsec [9] encryption and Mobile IPv6 [8] do not combine well to hide the mobile users identity. The reason is a kind of chicken and egg problem: it is not clear whether IPsec should run on top of Mobile IPv6 or vice versa. On one hand, the messages exchanged by the mobile node and its correspondents contain the mobiles home address, which identifies the mobile, in the home address option or routing header. Therefore, the headers of these messages should be encrypted to hide the home address. On the other hand, IPsec identifies endpoints based on their IP addresses. Therefore, it would be technically more elegant to run IPsec on top of Mobile IPv6 so that the IPsec policies and security associations can be indexed by the permanent home address rather than by the changing care-of address.

One solution, although not an ideal one, is to create a new security association between the correspondent and each new care-of address. The encryption would hide not only the home address but also the fact that a mobility protocol is being used. The tunnel would be encrypted but it need not be authenticated. Indeed, authentication is not even possible below the Mobile IPv6 layer because the only available identifier for the mobile is the care-of address. The session key for the encrypted tunnel between the care-of address and the correspondent could be created using an opportunistic key exchange, or with a unidirectional authentication and key exchange protocol.

An opportunistic key exchange is one that does not bind the session key to any identifier that has meaning on the application layer but, instead, uses whatever (even weak) credentials are available at the IP layer to prevent man-in-the-middle attacks. Encryption based on such a key exchange is useful for privacy protection of the content data and header data above the IPsec layer (including the care-of address) because it increases the cost of privacy violations.

A unidirectional authentication and key exchange, on the other hand, assures the mobile about the identity of the correspondent but gives the correspondent no information about the mobile. The asymmetry is similar to the way the SSL protocol is usually used. Since the goal is to protect the mobile's identity, the asymmetric authentication is sufficient. A limitation of both the opportunistic and the unidirectional protocols is that the resulting session key must not be relied upon for strong authentication of encrypted application-layer data.

This kind of excessive layering of IPsec is, of course, not an ideal solution. The fact that we have to resort to such trickery calls into question the whole Mobile IPv6 and IPsec protocol architectures. Since there is such intricate interaction between the secure tunnel and the endpoint-mobility protocol, it would make sense to combine the two into one protocol. Indeed, mobility could be implemented by changing the addresses of the endpoints of IPsec tunnels, instead of the degenerate kind of tunnel that Mobile IPv6 implements.

The above idea appears in the novel HIP and MOBIKE protocols, but also in the standard Mobile IPv6. The IPsec security association between the mobile node and its home agent is updated dynamically when the mobile moves so that it can always be looked up by the mobile's current care-of address. This makes it possible to send binding updates and tunneled data to the home agent through an encrypted IPsec tunnel, which hides the home address and the fact that Mobile IPv6 is being used. Since the security association between the mobile and its home agent is typically configured manually or by some out-of-band protocol, there is no key exchange that could divulge the mobile's identity to outsiders. If a dynamic key exchange is needed, it is important to use a protocol, such as IKEv2, that can hide the endpoint identifiers from passive eavesdroppers. It is also important that the key exchange protocol does not insist on using the IP address as an endpoint identifier (as most IKEv1 implementations do).

4.4 RFC-3041 addresses and DHCP

Apart from the permanent Mobile IP home addresses, most IP addresses are assigned dynamically either by DHCP or stateless autoconfiguration (the latter only in IPv6). Nevertheless, all IP addresses enable observers to correlate packets sent by the same host as long as the host keeps using the same address. Moreover, typical IPv6 addresses have a structure that essentially includes an embedded permanent identifier.

The interface identifiers in IPv6 addresses were originally created by expanding the 48-bit link-layer (MAC) address of the network interface card to a 64-bit EUI-64 [7]. Anyone who sees such an address can recover the MAC address and, thus, identify the network interface card. The unique identifier embedded into

the IPv6 address can be used to track a mobile device and user even outside the access network where one would not expect the link-layer address to be visible.

A solution was proposed in RFC 3041 [11]: the interface identifier can be selected randomly instead of using the EUI-64, and the value can be changed periodically. This makes no difference to the functionality of the address as the only real requirement for the interface identifier is that it is different for each node in the network. It is very unlikely that two randomly generated addresses collide. RFC-3041 addresses are particularly attractive for roaming mobile hosts that move from one access network to another because they can create a new random address at each access network and, thus, hide the link between the addresses.

IPv4 addresses cannot be randomized in the same way. On the other hand, IPv4 addresses do not have any part that uniquely identifies the host. DHCP servers typically allocate the same address to the same host every time it asks for an address, unless it has already been given to someone else. One possibility is to change this behavior so that the DHCP server randomizes the address assignments. This prevents peer endpoints and observers outside the local network from correlating connections with earlier ones when the mobile returns to a previous access network.

It is interesting to note that IPv6 hosts can anonymize their addresses independently, without any help from the network infrastructure. In IPv4, it is the network that needs to randomize the address assignment. This difference is, of course, a direct consequence of the fact that the preferred method of assigning IPv6 addresses is stateless autoconfiguration, where the host selects its own address, while IPv4 addresses are usually assigned by a stateful DHCP server.

The randomized addresses only help when the mobile is communicating with remote nodes across the Internet. Nodes on the access network can observe the mobile hosts link-layer address, which is a permanent identifier for the network interface hardware. The obvious solution would be to randomize also the link-layer address. Unfortunately, there is no consensus for doing this. In fact, the trend is quite the opposite: since Ethernet MAC addresses are used as a weak form of host authentication in many wireless LANs, the hardware manufacturers and driver writers are making the changing of MAC addresses more difficult than it used to be.

4.5 NAT and IPv6

The main reasons for deploying a network address translator (NAT) are that it allows multiple Internet hosts to share a single IP address and that it acts as a kind of stateful firewall by preventing Internet nodes from sending unsolicited packets to the local network behind the NAT. The fact that the ISP cannot differentiate pricing based on the number of nodes on the local network is particularly important for home users. Such per-host charging would, for example, prevent users from taking mobile devices from work to roam at the home network.

From the privacy point of view, there is also the advantage the NAT hides the structure of the local network from the Internet. Observers on the Internet

cannot tell whether two connections from the NAT originate from the same host or from two different hosts. If the NAT hides a more complex intranet that is composed of multiple local links, the subnet prefixes and, thus, the topology of the intranet, are also masked by the NAT. In addition to preventing the ISP from charging per node, an advantage of the NAT for mobile users is that it makes it much more difficult for Internet nodes to observe the arrival and departure of mobile nodes to and from the intranet.

A disadvantage of a NAT is that it makes providing global services from behind the NAT cumbersome and, in some cases, impossible. For example, it is not possible to have two web servers at home using the standard port 80. Also, NATs prevent the deployment of mobility and security protocols, such as Mobile IP and IPsec, which are both either blocked by a NAT or require complicated NAT-traversal solutions.

IPv6 solves the problem of IP-address shortage, which makes it less acceptable for ISPs to differentiate pricing based on the number of IP addresses used in each network. Eventually, more users will want to have server computers at home, such as web and file servers. They will also want to use IPsec, VoIP, and other protocols that are currently blocked or made inconvenient by NATs. Thus, there will be a need to allocate an IPv6 subnet prefix to each home, rather than a single IP address.

Many network administrators are, however, so used to the security provided by a NAT that they find the deployment of NAT-less IPv6 to be a threat rather than a promise. If this is the case, it is worth remembering that a stateful firewall can implement the same policies as a NAT and provide additional flexibility, e.g., by authorizing any intranet node to act as a server. The privacy aspect of NATs is, on the other hand, not as easy to replicate in IPv6 without actually bringing back the limitations of a NAT.

The solution we propose is to use more addresses instead of less. That is, an IPv6 hosts should use large numbers of temporary addresses and change them frequently. In fact, we propose using a new IP address for each TCP connection that originates from an IPv6 host. Using a new random or pseudo-random address for each connection makes it more difficult for an observer on the Internet to correlate two connections from the same host. It also prevents the observer from counting the hosts in the local network. Note that this solution is more useful for home users and other small networks with only a single subnet prefix than it is for large networks with multiple prefixes.

In order to prevent the TCP port numbers from revealing the association between the different hosts, it is better to either use the same port (e.g. 12345) or a completely random local port number for all TCP connections initiated by an IPv6 host. For UDP-based protocols, a similar approach can be taken, although it is necessary to understand the higher-layer protocol or application to know when it is appropriate to generate a new IP address.

If IPv6 hosts routinely use a large number of addresses, that may also guide the ISP charging policies away from counting the number of addresses, which will help create an IPv6 Internet without the unnecessary complications caused

by NATs. This would make it much easier to deploy home servers and mobility and security protocols.

4.6 Random addresses and CGA

Some network-layer signaling protocols for IPv6 rely on cryptographically generated addresses (CGA) [1] for message authentication. CGAs are IPv6 addresses where some bits of the address have been generated by computing a one-way hash of the mobile hosts public signature key and some additional parameters, including a random modifier. This key is then used to sign signaling messages. CGA-based authentication is being standardized for neighbor discovery and there are proposals for using CGA for authentication of mobility signaling and for end-to-end key exchange.

The problem with a CGA is that the address and the public key may be used to identify the mobile. If the mobile uses the public key to authenticate end-to-end signaling such as Mobile IPv6 binding updates, the only solution is to change the public key. However, if the mobile does not use signatures for end-to-end signaling, it can randomize its address by re-randomizing the modifier parameters in the hash input. (For those familiar with the CGA details: the re-randomization prevents the mobile from using long hash extensions but that hardly matters for mobiles that move frequently.)

The public key, of course, can also reveal the mobiles identity to the nodes in the same local network. This only makes a difference if the mobile is already randomizing its link-layer address or changes network interface cards occasionally. In that case, the solution is also to replace the public key and to generate a new CGA whenever the mobile moves from one access network to another.

5 Conclusion

In this paper, we discuss the privacy implications of transparent wireless and mobile Internet access, with focus on information present in packet headers and in signaling messages. We discuss the ways in which user identity and location can be revealed by the network protocols, usually without the user being aware of the disclosure, and the ways in which various network elements can use communications metadata to control the users Internet access. We also suggest simple technical solutions that reduce the amount of information leaked in order to preserve the users privacy and to retain more control at the end nodes.

References

1. Tuomas Aura. Cryptographically generated addresses (CGA). In *Proc. 6th Information Security Conference (ISC'03)*, volume 2851 of *LNCS*, pages 29–43, Bristol, UK, October 2003. Springer.
2. David Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1(1):65–75, 1988.

3. Alberto Escudero-Pasqual. *Privacy in the next generation Internet: data protection in the context of the European Union*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, December 2002.
4. Digital cellular telecommunication system (phase 2); Security related network functions. ETSI TS 100 929, European Telecommunications Standards Institute, November 1999.
5. Hannes Federrath, Anja Jerichow, and Andreas Pfitzmann. MIXes in mobile communication systems: Location management with privacy. In *Proc. Information Hiding Workshop*, Cambridge, UK, 1996.
6. Marco Gruteser and Baik Hoh. On the anonymity of periodic location samples. In *Proc. Intl. Conference on Security in Pervasive Computing*, volume 3450 of *LNCIS*, pages 179–193, Boppard, Germany, April 2005. Springer.
7. Robert M. Hinden and Stephen E. Deering. IP version 6 addressing architecture. RFC 3513, IETF Network Working Group, April 2003.
8. David B. Johnson, Charles Perkins, and Jari Arkko. Mobility support in IPv6. RFC 3775, IETF Mobile IP Working Group, June 2004.
9. Stephen Kent and Randall Atkinson. Security architecture for the Internet Protocol. RFC 2401, IETF Network Working Group, November 1998.
10. Dogan Kesdogan, Hannes Federrath, Anja Jerichow, and Andreas Pfitzmann. Location management strategies increasing privacy in mobile communication. In *Proc. 12th International Information Security Conference*, pages 39–48, Samos, Greece, 21–24 1996. Chapman & Hall.
11. Thomas Narten and Richard Draves. Privacy extensions for stateless address autoconfiguration in IPv6. RFC 3041, IETF Network Working Group, January 2001.
12. Pekka Nikander, Jukka Ylitalo, and Jorma Wall. Integrating security, mobility, and multi-homing in a HIP way. In *Proc. Network and Distributed Systems Security Symposium (NDSS'03)*, pages 87–99, San Diego, CA USA, February 2003.
13. Charles Perkins. IP mobility support. RFC 2002, IETF, October 1996.
14. Jonathan Rosenberg, Henning Schulzrinne, Gonzalo Camarillo, Alan Johnston, Jon Peterson, Robert Sparks, Mark Handley, and Eve Schooler. SIP: Session initiation protocol. RFC 3261, IETF, June 2002.
15. Andrei Serjantov. *On the Anonymity of Anonymity Systems*. PhD thesis, University of Cambridge, Cambridge, United Kingdom, March 2004.
16. Pyda Srisuresh and Matt Holdrege. Ip network address translator (NAT) terminology and considerations. RFC 2663, IETF, August 1999.
17. Alf Zugenmaier. *Anonymity for Users of Mobile Devices through Location Addressing*. PhD thesis, University of Freiburg, 2003.