# European Union Data Protection Policy 'Location privacy in the next generation mobile Internet'

Alberto Escudero Pascual <aep@kth.se>
Royal Institute of Technology. KTH
IMIT - IT University
Kista. Stockholm. Sweden

24th April 2002

## Abstract

The global telecommunication infrastructure will slowly converge toward an integrated packet switched network using the Internet Protocol as the common communication technology. First evidence of this convergence is the deployment of the third generation wireless infrastructure that brings together the radio access network and core network by using the next generation Internet Protocol IPv6.

The paper presents how 'mobility' is supported in IPv6. Mobility is the capability of mobile terminal to be reachable by its home network IP address with independence of the point of attachment to the Internet. We show the kind of information items that are required to be in transit in the network to allow a mobile node to seamlessly communicate on the move.

The European Commission proposal for a Directive (COM(2000)385) on 'processing of personal data and protection of privacy in the electronic communication sector' is kept in a technological neutral manner which means that no standards are imposed. It includes definitions inter alia on 'location data' and 'traffic data' and foresees privacy safeguards and different levels of protection for distinct kind of data. After the technical and legal overview we discuss the difficulties to apply the definitions provided by the Directive to certain technology as mobility in IPv6.

Based on the reasoning included in the paper we argue that classifying and defining data by traditional means and ways without taking into account Internet's multi-layered architecture might lead to an insufficient level of privacy protection for certain sensitive data and might not be the most appropriate way to adapt and update the existing provisions to new and foreseeable developments in electronic communications services and technologies.

## Introduction

This paper is divided as follows: Section 1 contains a brief overview of the Internet Protocol Version 6 and how mobility is supported, Section 2 introduces the proposed European Union Directive COM(2000)385 concerning the processing of personal data and the protection of privacy in the electronic communication sector with special remark to the privacy protection of location and traffic data, Section 3 presents some of open issues when trying to apply the Directive in the context of IPv6 mobility concerning the interpretation of the 'technology-neutral' definitions of traffic, content and location data. Finally, Section 4 presents some conclusions and regulatory recommendations.

# 1 The next generation Internet

The Internet Protocol Version 6 (IPv6), also known as "IPng" (IP Next Generation), is the latest version of the Internet Protocol (IP). Formally, IPv6 is a set of specifications from the Internet Engineering Task Force (IETF). IPv6 is being designed as an evolutionary set of improvements to the current IP Version 4. The most obvious improvement in IPv6 over the IPv4 are that IP addresses are lengthened from 32 bits to 128 bits which anticipates the future growth of the Internet and provides relief for what was perceived as an impending shortage of network addresses. Besides, IPv6 offers technical advantages over IPv4, including self-configuration mechanisms, enhanced security, quality of service features and native mobility support [1]. IPv6 aims to be the protocol capable of bringing together access and core networks, the 'glue' for the deployment of the future 'all-IP' telecommunication network.

IPv6 includes a security protocol in the network layer that provides cryptographic security services that supports combinations of authentication, integrity, access control, and confidentiality. The IP Encapsulating Security Payload ($ESP$) and the IP Authentication Header ($AH$) are part of the IP Security architecture (IPSEC) described in RFC 2401[2].

Both ESP and AH are mandatory parts of IPv6 and make sure that a third party eavesdropping on the channel can not read and/or modify the IP datagram. The IP Authentication Header seeks to provide security by adding authentication information to each IP datagram whilst confidentiality requires the use of ESP. Neither AH nor ESP hide the source and destination IP addresses of the communicating parties and hence their network location.

## 1.1 Mobility support in IPv6

The protocol operation defined for mobility in IPv6 is known as MobileIPv6 [3] and allows a mobile node to move from one link to another without changing the mobile node's IP address. A mobile node is always addressable by its 'home address' (HoA), an IP address assigned to the mobile node within its home subnet, i.e., with the prefix of its home link.

MobileIP allows users to move between different networks, while maintaining an addressable static identifier (home address). This is done by associating a dynamic identifier (care-of-address) with the mobile node when it is away from home. All traffic to the mobile node is intercepted in the home network by a home agent (HA) that tunnels the data to the care-of-address that is in use in that moment. Packets may be routed to the mobile node using their home address regardless of the mobile node's current point of attachment to the Internet (CoA), and the mobile node may continue to communicate with other nodes after moving to a new link. With MobileIP the movement of a mobile node away from its home link is thus transparent to transport and higher-layer protocols and applications.

MobileIPv6 shares many features with MobileIPv4[4], but the protocol is now fully integrated into IPv6. MobileIPv6 works in a similar way as MobileIPv4 does when using mobile node co-located care of address mode and route optimization. As in MobileIPv4 the mobile mode is responsible for discovering its current location. When the mobile mode is attached to its home link it directly receive packets and when roaming in a foreign network, it must acquire a co-located care of address and notify its home agent of this address.

MobileIPv6 on the other hand also includes the Mobility Header a new IPv6 protocol that allows the mobile node to inform selected IPv6 correspondent hosts of its care-of-address, so packets from these correspondent hosts can be redirected straight to the mobile node instead of using the home agent as an intermediary. The association between a mobile node's home address and care-of address is known as a 'binding' for the mobile node. A mobile node typically acquires its care-of address through stateless or stateful (e.g., DHCPv6) address autoconfiguration and while away from home registers its care-of address with a router on its home link, requesting this router to function as his home agent.

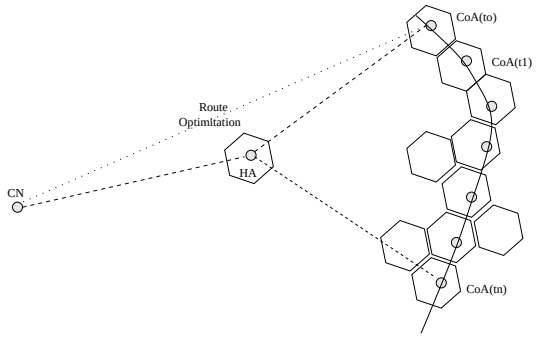


Figure 1: MobileIPv6. Node reachability

### 1.1.1 Binding dynamic (CoA) and static identifiers (HoA)

This binding registration is done by the mobile node sending to the home agent a packet with a Mobility Header containing a Binding Update message; the home agent then replies to the mobile node by returning a packet containing a Binding Acknowledgment message.

The Mobility Header by a set of different Binding (Request, Acknowledgment, Update and Missing) messages allows correspondent hosts communicating with a mobile node, to dynamically learn and cache the mobile node's binding (HoA-CoA).

Before sending a packet to any IPv6 destination, a node checks its cached bindings for an entry for the packet's destination address. If a cached binding for this destination address is found, the node uses an IPv6 Routing header to route the packet to the mobile node by way of the care-of address indicated in the cached binding. No IPv6 encapsulation is required due to IPv6's routing header. If a mobile node is at home destination the mobile node sees the home address in the routing header and processes the packet. If the sending node has no cached binding for the destination address, the node sends the packet normally to the home address without any Routing header and the packet is subsequently intercepted and tunneled by the mobile node's home agent. The correspondent nodes can also send a Binding Request message to the mobile node home address in order to obtain its care-of-address.

In summary mobility is enabled in IPv6 by two basic functions:

- **Mobility Header:** A mechanism to keep informed to the home agent and correspondent nodes of the changes of the mobile node's point-of-attachment to the Internet.

- **Efficient Routing:** Enable by including the home address in the destination option of *each* of the packets send by the mobile node to the correspondent node and by including the mobile node's home address in the routing header when the correspondent node talk to the mobile node via its care-of-address.

In the first case the binding process makes use of certain protocol named Mobility Header while the efficient routing is achieved by including the home address of the mobile node *as part of each* of the packets exchanged.

# 2 EU Data Protection Directive

On July 12, 2000 the European Commission adopted a proposal for a Directive (COM(2000)385) on 'processing of personal data and protection of privacy in the electronic communication sector'. The proposal is part of a package of proposals for initiatives which will form the future regulatory framework for electronic communications networks and services. It aims to adapt and update the existing Data Protection Telecommunications Directive (97/66/EC) to take account of technological developments.

Unlike the previous Telecommunications Directive

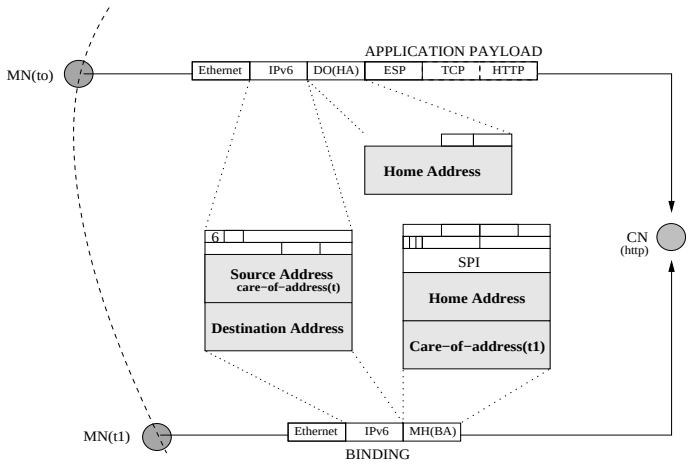


Figure 2: Mobility/Location Information embedded in IPv6 headers

the scope of the future Directive would not be restricted to telephony and data networks, but will also cover satellite, terrestrial and cable TV broadcasting networks, irrespective of the type of information concerned.

The new proposed EU Data Protection Directive establishes a common framework for data protection in telecommunication services and networks regardless of the technology in use in electronic communication services and networks. The Directive provides a set of protections or safeguards and various definitions such as: communication channel, traffic and location data, service provider, etc.

As part of the proposed changes, the new Directive replaces the existing definitions of telecommunications services and networks in Directive (97/66/EC) to align the terminology with a common framework for electronic communications services and networks. The changes of definitions try to ensure that all kind of electronic communications will be covered regardless of the technology in used. The Directive also include four new important definitions to strengthen the common understanding. They relate to the following:

- **Communication** is any information exchanged or transmitted between a finite number of parties by means of a publicly available electronic communications service.
- **Call** is a connection established by means of a publicly available telephone service allowing two-way communication in real time.
- **Traffic Data** is any data generated and processed in the course of or for the purpose of the transmission of a communication over an electronic communications network.
- **Location Data** is any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service.

For the purpose of this paper we consider the protections related to traffic and location data. The Article 6 of the directive prohibits the use of traffic data expect for billing purposes and introduces the possibility for further data processing for value-added services based on consent of user/subscriber. Article 9 introduces specific privacy safeguards for subscribers and users with regard to mobile location information services.

# 3 Open issues in mobile privacy

In this section we present two issues that arises when trying to apply the Directive to the concrete scenario of mobility in IPv6. The first one concerns the distinction between traffic and content of the communication in the Internet and the second is related to the meaning of geographical position.

## 3.1 From traffic data to content data

Traditionally content data is considered to be more sensitive than traffic data and therefore embedded in a higher level of privacy protection. But on the Internet content and traffic data can only be clearly distinguished when given a very concrete context, item of interest and level of observation [8]. Internet Protocols take benefit of a multi-layered architecture, where higher-level components talks to lower-level components and vice-versa. The result of this layered architecture is that what can be seen as traffic/signaling for one layer can be indeed be content for a lower layer.

The benefits of a layered architecture is that the communications functions are partitioned into a set of hierarchical layers where each layer performs a related subset of the functions required to communicate relying on the next lower layer to perform more primitive functions while provides services to the next higher layer.

But it is not only the Internet's multi-layered architecture which makes a clear distinction between content and traffic data difficult. Moreover, the issue of interest plays a fundamental role in the discussion, what can be pure traffic or signaling data for one observer can be considered content for another. For example, the fact that two parties are communicating in a given time can provide an observer with a certain amount of sensitive information not requiring to examine the 'formal content' of the communication at all.

Another clear example of this complexity can be found in mobility in IPv6. As shown in [1.1] MobileIPv6 supports transparency above the IP layer that includes the maintenance of active TCP connections when a mobile node changes its point of attachment to the Internet. From the point of view of the application, the mobile node always use the mobile node's home address as being statically in the home network.

The binding messages can be seen as pure 'mobility signaling' hidden to the application or on the contrary as a rich content information that allows to the mobile node to make location aware decisions.

In its provisions the Directive tries to adapt the traditional mechanism to differentiate signaling and content in traditional plain old telephone systems (POTS) to the Internet or other packet switched networks and services. In POTS signals and data are transported in different channels and include a very restricted number of message formats. The signaling channels take care of the interactions between the components. Typically, this information is either data being transferred between end-users and other end-users (known as content) or between end-users and the network (traffic).

The traditional way of classifying the data conflicts with the pure nature of a multi-layered architecture of the Internet where the distinction depends on the functionality and layer of observation. What can be seen as one communication process from the point of view of the user (exchange of e-mail between two mail agents) involves multiple levels of signaling and content.

## 3.2 From traffic data to geographical position

The second open question is related to definition of geographical position and the Internet addressing.

As shown in [1.1.1] the mobility header and the efficient routing make sure that the communicating parties are aware of the changes of the mobile node attachment to the Internet. Opposite to the old communication systems, where signaling and content are transmitted in certain channels, what could be considered as 'IPv6 mobility signaling' is a set of IP packets or/and packet's fields. [Fig.2] shows the structure of two 'mobileipv6' packets exchanged between a mobile node and a web server, the first one is sent during the web browsing session and the second is a 'binding update'. The figure also shows how mobility related

information is present in both packets and embedded in two different ways in the IPv6 header.

When an eavesdropper is located somewhere along the route between the home agent and the mobile node it is possible to identify and track the mobile node movements by observing the 'mobility signaling' exchanged between the different communicating parties. The mobility of a certain mobile node [Fig.1] with home address $HoA_i$ in a period of time $(t_o < t < t_n)$ can be represented as a series of care-of-addresses $[CoA(t_o), CoA(t_1), CoA(t_2)...CoA(t_n)]_i$.

It is out of the scope of this paper to describe the different Geographic Information Reference Systems but while the care-of-address doesn't represent a standardized geographical position due to the nature of the reference system (the Internet infrastructure), the care-of-address and its changes during the time can without any doubt reveal mobility and if not absolute, relative positioning [7].

Part of the ongoing research work of the *PETng Project [6]* is to create a model that will allow to an eavesdropper to determine the proximity of two mobile nodes by observing the changes of their CoA(t) during the time i.e. to determine the 'relative positioning' of two mobile nodes.

# 4 Conclusions

Traditional legal, regulatory and technical provisions were established with traditional technological environments in mind. When telephone traffic data was decided to be less invasive than the content of the conversation, this reflected the plain-old-telephone system (POTS): traffic data was merely the person who was calling or the person called, and the duration. Accordingly, one level of privacy protection was typically assigned to traffic data (if at all), and another was applied to access to communications content, that is, the conversation itself.

The traditional classification of data based on which functional channel is sent and received conflicts with the pure nature of a multi-layered architecture of the Internet where the distinction depends on the layer of observation and the interactions that are observed.

In order to illustrate the complexity of using technology-neutral language we have shown how problematic is to classify the data in traffic, content and location in mobile Internet. What it can be considered as traffic data can be either content or location depending on the items of interest for the observer and/or application.

It is our concern that technology-neutral language may be used to ignore, willful or not, the challenges, risks, and costs to applying powers to different technical infrastructures. Classifying and defining data by traditional means and ways without taking into account Internet's multi-layered architecture might lead to an insufficient level of privacy protection for certain sensitive data based on the fact that they are grouped under certain category.

If policy makers insist on applying traditional powers to these new infrastructures, we argue that the new lawful policies must acknowledge that the data being collected now is separate from tradition and can only be understood having the different technologies in mind.

# 5 Acknowledgments

I will like to acknowledge to Ian (Gus) Hosein from the Department of Information Systems at the London School of Economics with whom i started to research in this interdisciplinary area and Corinna Schulze from the Directorate General Information Society of the European Commission for giving background information on the existing legal Framework on Data Protection.

Lastly i want to express my gratitude to the Personal Computing and Communication (PCC) research program and the Swedish Center for Internetworking (SCINT) for supporting my work in Internet privacy.

# References

[1] **C. Huitema,** *IPv6, the new Internet Protocol.* 2nd Edition. Prentice Hall. 1997.

[2] **S. Kent and R. Atkinson**, *"Security Architecture for the Internet Protocol".* RFC 2041.

[3] **D. Johnson and C. Perkins,** *"Mobility Support in IPv6"*, draft-ietf-mobileip-ipv6 v16. March 2002.

[4] **C. Perkins,** *"IP Mobility Support"*, RFC 2002, October 1996.

[5] **A. Escudero**, *"Anonymous and Untraceable Communications: Location Privacy in Mobile Internetworking".* Licentiate Thesis. July 2001.

[6] **PETng Project.** *"Privacy Enhanced Technologies in the next generation Internet".* SCINT/KTH/KAU. http://www.petng.net

[7] **A. Escudero**, Contribution to the EU Forum on cybercrime. Location data and traffic data. Brussels. November 2001.

[8] **I. Hosein, A. Escudero,** *"Understanding traffic data and deconstruction technology-neutral regulations".* UNECE. March 2002.