Anonymity for Users of Ubiquitous Computing

— extended abstract on work in progress —

Alf Zugenmaier[1], Adolf Hohl[2]
[1]Microsoft Research
Cambridge, UK
alfz@microsoft.com

[2]Institute for Computer Science and Social Studies, Department of Telematics
Albert-Ludwigs-University
Freiburg, Germany
ahohl@iig.uni-freiburg.de

***Abstract*–Anonymity is a protection goal that helps to protect the privacy of users by ensuring that their identity remains unknown. As privacy is a grave concern in pervasive computing, the need for suitable anonymity mechanisms is apparent. This contribution uses the Freiburg privacy diamond to analyze the possibilities for anonymity mechanisms in pervasive and highlights the problems that arise out of the one user many devices model.**

I. Introduction

The paradigm of ubiquitous or pervasive computing [1] leads to a much greater intrusion of information and communication technology into the personal life of everyone than what we experience today. The users of pervasive computing will use many smart personal objects, in addition many services will be provided by the smart environment that is envisioned to surround us. As the user moves around, his personal devices move around as well. From the communications perspective, therefore, the types of mobility to consider are user and device mobility.

In a world with ubiquitous computing (UC), privacy becomes more important for the users because there isn't any more a "private zone" into which the user can retreat [2][3]. Important parts of the user's privacy in this context are: privacy of the user's location, privacy of the user's actions and privacy of the user's identity. In this paper we focus on the latter, i.e. anonymity, in communicating with the UC applications.

"Anonymity" is described in the Oxford English Dictionary as "the state of being anonymous, used of an author or his writings", with anonymous meaning "nameless, having no name". The standardization body ISO defines in the Common Criteria [5]: "...[Anonymity] ensures that a user may use a resource or service without disclosing the user's identity."

The definition of anonymity that is adopted in this paper is following: An action fulfils the condition "*anonymity*", if the attacker is not able to deduce the identity of the originator, also called user, from that action or its context. The identity is the set of all personal data and uniquely identifying subsets thereof [10],[11]. An action[1] is anything the user does which takes a limited time, i.e. has a defined start and end time and has an instantaneous effect. The context of an action is all additional knowledge an attacker can gain about this action – it includes the request and reply messages, the log files of the server, etc. An action can be as granular as sending a single message, or contain a complete transaction. If pseudonyms are used,

Part of this work for this presentation was funded by the Gottlieb Daimler and Carl Benz Foundation. Part of this work was done within the EU project WiTness.

[1] Action is defined by the Oxford English Dictionary as a thing done, a deed.

the user is also considered to be anonymous for as long as the assumed attacker is not able to reveal the identity of the user using the pseudonym.

Information about the user of a service can be gathered from the content of the messages between service and user, the source and destination addresses of these messages, and the traffic flow. To protect content information from an attacker in the communication system, the messages can be encrypted. If that is not possible, omitting all references to the user in the content of the messages serves to protect the user's identity from the communication partner and everyone else who is able to access the content of the message.

It is more difficult to conceal the source and destination addresses of the messages. If a broadcast communication is possible, as in television, the intended recipient of a message can be concealed. Nevertheless, the sender of a message can still be determined, in this case the broadcasting company. Absolute sender anonymity can be achieved with DC-networks [12],[13].

For networks that do not have the broadcast property a number of anonymizing techniques exist at the application and network level. An overview is provided, e.g., in [14]. The simplest anonymizing technique is to use a relay, also called a proxy, for requests and replies. The relay removes identifying information from the headers of the request. This way the web server does not receive the address of the identity of the user; the user is anonymous to the communication system apart from the relay. But the communication is still vulnerable to traffic analysis.

To overcome the limitations of proxy based solutions, it is possible to use the Mix concept [15]. Several mixes are integrated into the network and messages are routed through these mixes. Every mix collects messages, waits for a certain number of messages to arrive, shuffles them, and then sends them to their next destination. Variants of this concept are Onion Routing [16], SG mixes [17], and Web-mixes [18]. They try to overcome performance limitations. To some extend the system “Crowds” [19] can also be seen as a variant of this concept.

In some mechanisms, as proposed by RFC 3041 [20] or the IEEE [21], the address of the user's device is replaced by a temporary address, therefore making it difficult to correlate the address of the device with the identity of the user. Location addressing [24] works in a similar manner, all references that can be used to identify a device are removed from the communication, on all network layers of the protocol stack.

In addition to these technical mechanisms there are "classical" anonymizing techniques such as using public pay phones or internet cafes.

We present how the Freiburg privacy diamond (FPD) [4], a model for analyzing anonymity mechanisms, can be extended to describe the additional challenges faced when designing an anonymity mechanism for pervasive computing. This model shows clearly, how all of the user's devices must work together to achieve anonymity. It also shows the limitations that one must impose on the sensors and devices in the user's environment. The extended FPD can be used to evaluate the usefulness of proposed anonymity mechanisms, as well as help in constructing new ones.

This paper is structured as follows: The next section describes the Freiburg privacy diamond. Section Three extends this concept to account for the fact that a user may use several devices. An example is given in which the communication using each device by itself is anonymous, but the combination leads to a breach of anonymity. A conclusion and an outlook are presented in Section Four.

II. Freiburg Privacy Diamond

The Freiburg privacy diamond (FPD) [4] is a model that tries to capture the essence of anonymity with regard to the most important forms of mobility in mobile communications: terminal mobility and user mobility. Therefore, it must consider at least four types of entities: the action itself, the terminal, i.e. the device used for the action, the user who performs the action and the location which the device and the user are located at.

The FPD (c.f. Figure 1) describes how these entities are related and how an attacker can use knowledge about these relationships to break anonymity. With this completely interconnected graph[1] it is possible to describe which information can be concluded from other information. The use of the FPD is illustrated in a

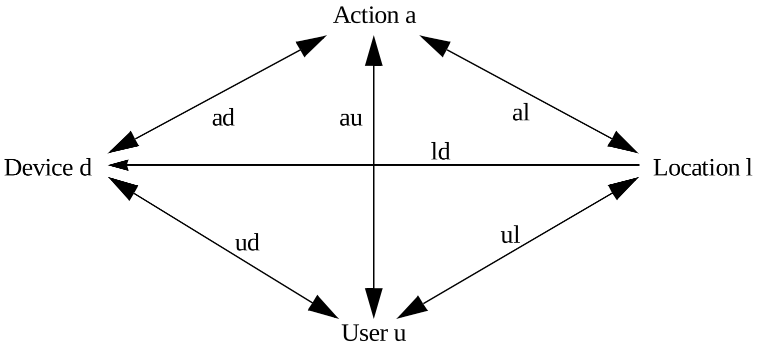

Fig. 1. The Freiburg privacy diamond

very simplified fashion by the following example. An attacker attempting to disclose the identity of a user tries to reveal the relationship between the user and an action. To do this, he could find out which device was used for this action and then find out who used this device. If the identity of the device used for the transaction is concealed, e.g. using a mix network, this deduction is not possible. Other conclusions, e.g. based on taking into consideration the location from which the action was carried out may, however, still be possible.

A. Classes of Anonymity Mechanisms

Intuitively, there are five loop-free paths which can be used to deduce the identity of a user by linking this action to the user:

1 user to action directly
2 user via location to action
3 user via device to action
4 user via location and then device to action
5 user via device and then location to action

For anonymizing systems that are secure against an attacker calculating the transitive closure, all five paths have to be broken. There are four minimal ways of doing this (cf. Figure 2), leading to four classes of minimal anonymity mechanisms. Minimal means that it is not possible to re-connect a severed relation in the privacy diamond without allowing the attacker to infer the relation of user to action through transitive closure.

Anonymizing mechanisms in the category described by the privacy diamond of Figure 2a are those that do not require mobility, e.g Mixes and DC-nets. The privacy diamond of Figure 2b describes anonymizing mechanisms that rely on user mobility like phone booths or Internet cafés. An anonymizing mechanism in category c) relies on broadcasts to and from a specific device. Both categories b) and c) rely on the users changing their devices. Therefore, it is not possible to employ a personal device. Category d) requires terminal mobility, enabling users to use their own devices. It also permits the location of the action to be visible to the network, thus allowing optimization of routing, etc. RFC3041 and location addressing are examples of mechanisms in that class.

It is possible for an anonymity mechanism to be contained in two classes if it is not minimal. This can happen by combination of mechanisms of different classes.

[1] The description of the FPD here is slightly different from the original as described in [4] to simplify the model while retaining all desired properties. The main difference is that an undirected rather than a directed graph is considered.

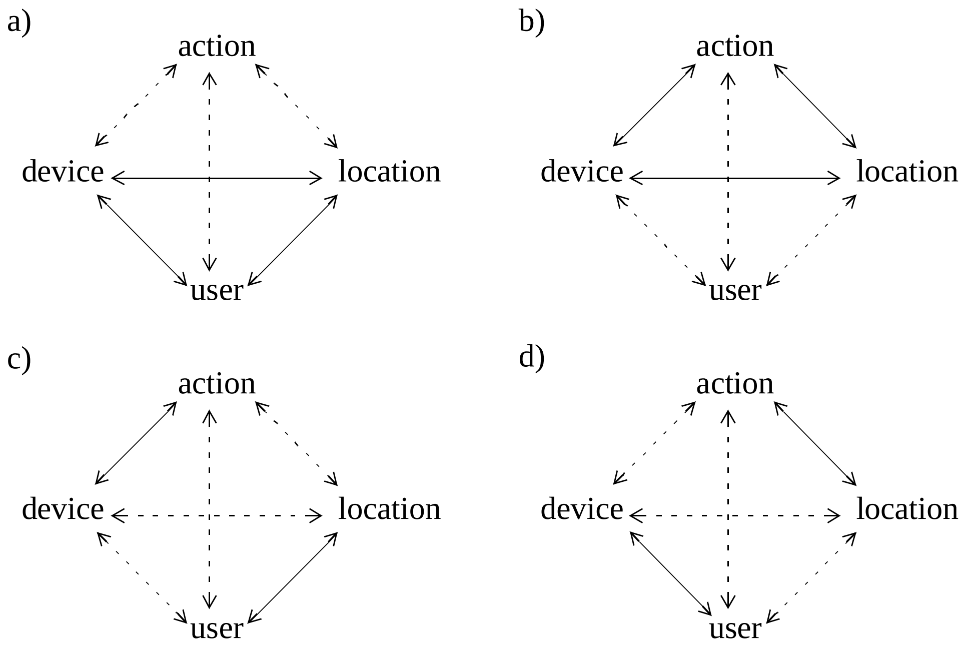

Fig. 2. Four minimal possibilities for anonymity mechanisms. Relations that must be obscured by the anonymity mechanism are shown by dotted arrows.

B. Model Assumptions

The privacy diamond is used to represent the knowledge of the attacker in the following situation. A user operates a device at a certain location to initiate an action. Four entities are used to model the situation: the user, the action, the location, and the device. Time will only be considered as an implicit parameter.

Mobile users use a mobile device to perform actions. These actions are considered to be short; during an action neither the user, the device of the user, nor the location of the user changes. The action is also instantaneous, since it is carried out while the user uses the device. One action at a higher layer can be made up of several actions on lower layers.

To model location information, the world is divided into cells. The size of these cells determines the maximum resolution with which a device or user can be located. This simplification reduces the mathematical complexity associated with continuous variables.

Users perform actions using a single device from a set of possible devices. The device is located at the same place as the user. This assumption is realistic, as the user has to be in the proximity of the device to operate it.

In this paper the scope of definition of the device includes all software on this device. If the software is able to migrate from device to device, as is the case with mobile agents, this node of the graph would have to be replaced by several nodes. This situation is not considered in this paper.

C. Semantics for the Privacy Diamond

The sets of entities involved have to be defined, as well as possible relationships and semantics for these relationships.

The following sets are defined for the privacy diamond:

$$A = \text{Set of all actions} \,, D = \text{Set of all devices}$$
$$L = \text{Set of all spatial cells} \,, U = \text{Set of all users}$$

The attacker has knowledge about the existence of finite subsets of these sets, which are: $A \subseteq A$, $D \subseteq D$, $L \subseteq L$, and $U \subseteq U$.

The semantics for the relationships is that the attacker has knowledge about:

ad which action could have been initiated from the device
al from where (location) the action could have been initiated
etc.

The answers to these questions can be of three types: a positive answer, stating which entities have a relationship; a negative answer, stating which entities do not have a relationship; and an indifferent answer, stating that nothing is known about the relationship between the entities.

D. The Privacy Diamond using Relations

When using symmetric relations to represent the knowledge of the attacker, the complexity of the model is reduced by considering only positive answers to the questions. The privacy diamond is formalized with 12 relations, $R_{XY} \subseteq X \times Y$, with $X, Y \in \{A, D, L, U\}$ and $X \neq Y$, i.e. $R_{AD} \subseteq A \times D, ..., R_{UL} \subseteq U \times L$. The symmetry implies that $R_{XY} = R_{YX}$. The statement for the relation R_{AD} on arrow AD means: which device d could have initiated action a? If (a,d) is contained in the relation R_{AD}, this means that device d could have initiated action a. This relation does not exclude the possibility that another device d' could also have initiated this action. If d or d' could have initiated action a, the relation R_{AD} would be $R_{AD} = \{(a, d), (a, d')\}$.

Definition: An element (x, y) is in a relation R_{XY} if the attacker has evidence suggesting that the entity x could be the answer to the question regarding which element is in a relationship with y. The nature of the relationship is in accordance with Subsection C.

The relation R_{AU} is the most interesting for an attacker attacking the user’s anonymity. The information about who could have performed the action is provided in this relation. In the best case, the attacker has no idea about who could have implemented the action; the relation contains no element linking the actions performed to a user. It could also be that the user conceals himself within an anonymity set; the relation contains many elements. In this case, the number of users who could have performed the action has to be greater than a threshold.

Definition “anonymity of an action”: an action, a, fulfils the condition of anonymity if the relation representing the knowledge of an attacker of who could have performed the action, R_{AU}, contains no element linking a to any user u, i.e.:

$$R_{AU} \cap (\{a\} \times U) = \varnothing \tag{1}$$

or if the user is able to conceal himself in an anonymity set that is greater than his personal threshold $t > 0$:

$$\left|R_{AU} \cap (\{a\} \times U)\right| > t \tag{2}$$

Definition “transitivity rule”: combining the knowledge represented by two relations, $R_{XY} \subset X \times Y$ and $R_{YZ} \subset Y \times Z$, where $X, Y, Z \in \{A, D, L, U\}$ and $X \neq Z$, new pairs of relations $R_{XZ} \subset X \times Z$ can be constructed following a “transitivity” rule, a standard relational composition:

$$R_{XZ, new} = R_{XZ} \cup R_{YZ}R_{XY} = R_{ZX, new} \tag{3}$$

observing the symmetry of the relation.

For example, if we know that device d initiated action a and that user u used device d at the same time, i.e. $R_{AD} = \{(a, d)\}$ and $R_{UD} = \{(u, d)\}$, it follows that $R_{AU} = \{(a, u)\}$, that is, user u initiated action a.

Definition: The union of the initial set of relations that are known to an observer O defines the observer’s initial view.

$$\mathrm{View}_O = R_{AD} \cup \ldots \cup R_{DL} = \bigcup_{X, Y \in \{A, D, L, U\}; X \neq Y} R_{XY} \tag{4}$$

Definition: The closure of View_O under the transitivity rule will be denoted by $\overline{\mathrm{View}_O}$.

$$\overline{\mathrm{View}_O} = \mathrm{View}_O \cup (\mathrm{View}_O^2 \cap S) \cup (\mathrm{View}_O(\mathrm{View}_O^2 \cap S) \cap S) \cup \ldots \tag{5}$$

The notation View_O^2 is used for the composition of View_O with itself. The set S,

$$S = D \times A \cup A \times D \cup \ldots \cup L \times D \cup D \times L \tag{6}$$

is introduced to eliminate the relations for which no semantics is defined. Because the sets A, D, L, and U are finite, only a finite number of relations are possible in View_O. Therefore, $\overline{\mathrm{View}_O}$ actually is the union of a finite number of sets only.

In addition to this attack which only makes use of information contained in one FPD, [23] describes three attacks which make use of context information. In this model, these attacks are described in terms of the attacker’s view before the attack $\mathrm{View}_{O,\ \mathrm{before}}$ and after the attack $\mathrm{View}_{O,\ \mathrm{after}}$. Likewise, these attacks can be performed on $\overline{\mathrm{View}_{O,\ \mathrm{before}}}$, i.e. after the transitive closure is calculated.

III. Extended Privacy Diamond

The presented tool for analyzing the anonymity of a user with a mobile device is able to model exactly one user with exactly one device. In a world of ubiquitous computing, it can be assumed, that users have more devices on them which are communicating among each other and with the environment. To analyze the anonymity of a user using a number of devices, every device has to be analyzed with a independent Freiburg Privacy Diamond (in a parallel process). This is visualized in the following figure, where four layers/devices of a privacy diamond are representing the relations of mobile devices to a user. For reasons of visibility, the relations R_{AU} and R_{DL} are not visualized in the figure, but are also part of the model.

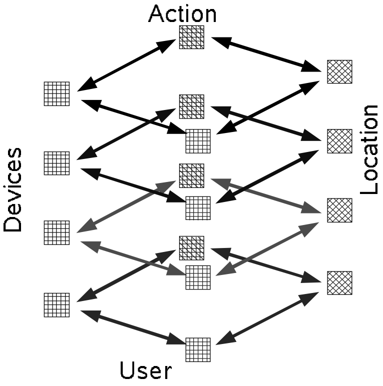

Fig. 3. A stack of Freiburg privacy diamonds

All these devices are operated by the same user, namely the one who wants to protect his or her anonymity, therefore the nodes of the user can be identified with one single node. Similarly, all devices are at the same location, so the nodes of the location can be simplified, too. A simplification of the action nodes cannot be made without exceptions, because not all devices are involved when an action is performed. However, often it is possible to identify multiple actions with a strong correlation of content or time with one action at a higher layer by means of grouping the actions together. These simplifications are visualized in Figure 4.

If a device is not used during performing an action, no relation R_{AD}, R_{AL} or R_{AU} exists for this layer/device.

A. Extended Definitions

In analogy to the definition ***"anonymity of an action"*** a user can remain anonymous if there is no relation from the user to the corresponding action. This means that the property of being anonymous can be expressed by the inexistence of a relation from the user to an action. It must be noted, that the ***"transitivity***

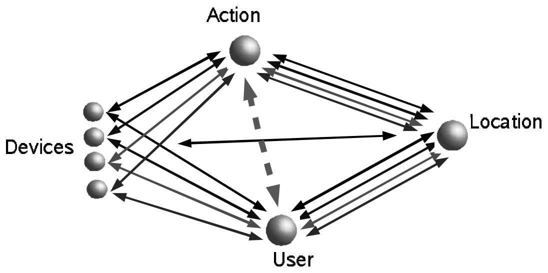

Fig. 4. Simplified extended privacy diamond

rule" has to be extended because a hop to a node on another layer is possible, if knowledge is available between these two nodes. The layer of the privacy diamond respectively the index of the user device is denoted as *i*. Now the knowledge of the attacker consists relations from $12 \times i_{max}$ sets of pairs, with i_{max} being the number of layers.

Extended definition "anonymity of an action": an action, a, fulfils the condition of anonymity if the relation representing the knowledge of an attacker of who could have performed the action, R_{AU_i}, contains no element linking a to any user u, i.e.:

$$R_{AU_i} \cap (U \times \{a\}) = \varnothing \tag{7}$$

The ***"transitivity rule"*** has to take into account the layers of the stacked privacy diamonds. If additional knowledge between two ore more layers is available, the transitivity closure can be deduced on the union set of the relations of these layers.

Extended definition "transitivity rule": combining the knowledge represented by two relations, $R_{XY} \subset X \times Y$ and $R_{YZ} \subset Y \times Z$, where $X, Y, Z \in \{A_i, D_j, L_k, U_l\}$ with $1 \le i, j, k, l \le i_{max}$ and $X \neq Z$, new pairs of relations $R_{XZ} \subset X \times Z$ can be constructed:

$$R_{XZ, new} = R_{XZ} \cup R_{YZ}R_{XY} = R_{ZX, new} \tag{8}$$

The transitive closure is extended in a similar manner by repeatedly applying the transitivity rule.

B. Examples.

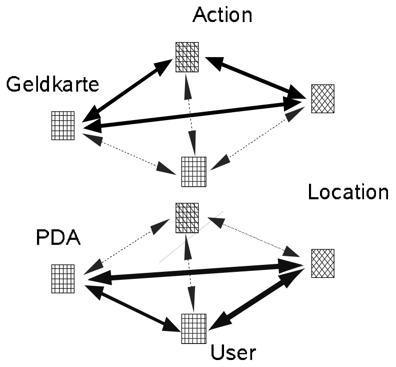

Fig. 5. Freiburg privacy diamonds of two devices

The first example illustrates the use of multiple devices: A user uses location addressing for his PDA to remain anonymous, i.e. only the relation between user and device and between location and action are known. He also owns a RFID cardkey. Whenever the cardkey is used, the location of the user becomes known. This leads to a compromise of the anonymity of the actions that the user performs.

A second example illustrates the use of incompatible anonymity mechanisms: A person owns two devices, which can be used preserving anonymity when they are used seperately. The first device is a personal digital assistant that can be used to access the Internet through access points at the bus stations for users with registered devices. The second device is an anonymous payment card like the Geldkarte in Germany[1]. The attacker does not know which person owns this particular payment card. The person waits at a bus stop with an access point. The user uses his PDA to browse the Internet using a Mix network. Then the bus arrives and is boarded. The person uses the payment card to pay for the ticket. Because the attacker may have the context knowledge of only one person waiting at this bus stop – this information may come from a sensor network – it is possible to combine the two privacy diamonds. Then the transitive closure of the extended FPD gives a relation between the user and the action breaking the anonymity of the payment. The relations in the FPDs from the use of these two devices and the resulting transitive relation is illustrated in Fig. 5 The thin dashed line represents relations unknown to the attacker. Fig. 6 shows all relations in the extended privacy diamond, where a path from the user to the action is available and the transitive closure can be deduced to find a way from the user to the action.

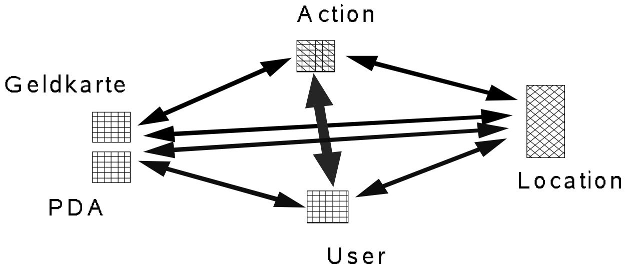

Fig. 6. Relations in the extended privacy diamond

IV. Conclusions

This paper introduced the Freiburg privacy diamond as a model which can be used to understand the impact of user and terminal mobility on anonymizing systems. Based on this model, an extended privacy diamond was introduced which can be used to analyse if a user is anonymous when using more than one device per user. The deduced results of the single Freiburg privacy diamond are transferred to the extended privacy diamond and an example shows, that the use of a combination of devices using incompatible anonymizing mechanisms can compromise the anonymity, which is achieved when each device is used seperately.

The concept of the privacy diamond only helps to understand anonymity in a communications context. Future work tries to transfer the method of using relations to represent the attackers knowledge to other areas, e.g. evaluating the content of the data an attacker has collected over the course of multiple transactions.

References

[1] Mark Weiser. The Computer for the 21st Century, pp 66-75, Scientific American, January 1991.

[2] Marc Langheinrich. Privacy by Design - Principles of Privacy-Aware Ubiquitous Systems. Ubicomp 2001. Springer-Verlag LNCS 2201, pp. 273-291, 2001.

[3] Marc Langheinrich. A Privacy Awareness System for Ubiquitous Computing Environments. Ubicomp 2002 Conference Proceedings. Springer-Verlag LNCS 2498, pp. 237-245, September 2002.

[4] Alf Zugenmaier. The Freiburg Privacy Diamond – A Conceptual Model for Mobility in Anonymity Systems. To appear in: *Proceedings of Globecom 2003*, 2003.

[5] Common Criteria for Information Technology Security Evaluation. *Part 2: Security Functional Requirements*. Version 2.1, August 1999.

[6] K. Rannenberg, A. Pfitzmann, and G. Müller. "Sicherheit, insbesondere mehrseitige IT-Sicherheit," in: G. Müller and A.

[1] It is possible to get a Geldcard anonymously, but as it uses a shadow account to prevent fraud, one has to make sure it is always used anonymously!

Pfitzmann, editors, *Komponenten, Integration*, volume 1, *Mehrseitige Sicherheit in der Kommunikationstechnik*, pp 21-29. Addison Wesley Bonn, 1997.

[7] K. Rannenberg. *Zertifizierung mehrseitiger IT-Sicherheit: Kriterien und organisatorische Rahmenbedingungen*. Vieweg Wiesbaden, 1998.

[8] R. Clarke. "Identified, Anonymous and Pseudonymous Transactions: The Spectrum of Choice," *Proc. of the User Identification & Privacy Protection Conference*, September 1999.
Accessed at http://www.ana.edu.au/people/Roger.Clarke/DV/UIPP99.html, on March 1, 2002.

[9] A. Pfitzmann and M. Köhntopp. "Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology," H. Federrath (Ed.): *Designing Privacy Enhancing Technologies; Proc. Workshop on Design Issues in Anonymity and Unobservability*; LNCS 2009; pp 1-9, 2001.

[10] D. Gerd tom Markotten, U. Jendricke, and G. Müller. "Benutzbare Sicherheit - Der Identitätsmanager als universelles Sicherheitswerkzeug," G. Müller and M. Reichenbach, editors, *Sicherheitskonzepte für das Internet*, pp 135-146. Springer-Verlag Berlin, 2001.

[11] U. Jendricke and D. Gerd tom Markotten. "Usability meets Security - The Identity-Manager as your Personal Security Assistant for the Internet," *Proc. of the 16th Annual Computer Security Applications Conference*. pp 344-353, December 2000.

[12] D. Chaum. "The dining cryptographers problem: unconditional sender and recipient untraceability," *J. of Cryptology*, pp 65-75, 1 (1) 1988

[13] M. Waidner and B. Pfitzmann. "Unconditional Sender and Recipient Untraceability in spite of Active Attacks - Some Remarks," Fakultät für Informatik, Universität Karlsruhe, Interner Bericht 5/89, March 1989. Accessed at http://www.semper.org/sirene/publ/WaPf_89IB_DCandFailStop.ps.gz on May 24, 2002

[14] M. Borning, D. Kesdogan, and O. Spaniol. "Anonymität und Unbeobachtbarkeit im Internet (Anonymity and Untraceability in the Internet)," *it+ti Informationstechnik und Technische Informatik*, pp 246-253, 5 (43) 2001

[15] D. Chaum. "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," *Comm. of the ACM*, pp 84-88, 2 (24) 1981.

[16] D. M. Goldschlag, M. G. Reed, and P. F. Syverson. "Hiding Routing Information," R. Anderson, editor, *Information Hiding*, LLNCS 1174, pp 137-150. Springer-Verlag, May 1996.

[17] D. Kesdogan, J. Egner, and R. Büschkes. "Stop-and-Go-MIXes Providing Probabilistic Anonymity in an Open System," *Information Hiding 1998*, LNCS 1525, pp 83-98, Springer Heidelberg, 1998.

[18] O. Berthold, H. Federrath, and S. Köpsell. "Web MIXes: A System for Anonymous and Unobservable Internet Access," H. Federrath, editor, *Designing Privacy Enhancing Technologies*, LNCS 2009, pp 115-129, 2001.

[19] M. Reiter, A. Rubin. Crowds: "Anonymity for Web Transactions," *ACM Trans. on Information and Systems Security*, pp 66-92, 1 (1) 1998.

[20] T. Narten and R. Draves. "Privacy Extensions for Stateless Autoconfiguration in IPv6," RFC3041, January 2001. Accessed at http://www.ietf.org/rfc/rfc3041.txt on June 21, 2002

[21] P. Orava, H. Haverinen, J.-P. Honkanen, J. Edney. "Temporary MAC Addresses for Anonymity," submission to IEEE P802.11, http://grouper.ieee.org/groups/802/11/Documents/D2T251-300.html accessed on 28/02/2003

[22] F. Dupont. "RFC 3041 Considered Harmful," Internet draft, http://www.globecom.net/ietf/draft/draft-dupont-ipv6-rfc3041harmful-00.html accessed on 28/02/2003

[23] O. Berthold, H. Federrath, and M. Köhntopp. "Project „Anonymity and Unobservability in the Internet“," *Proc. of the Workshop on Freedom and Privacy by Design / CFP 2000*, 2000.

[24] Alf Zugenmaier. *Privacy for Users of Mobile Devices though Location Addressing*. Rhombos-Verlag, Berlin, 2003.