angenommen bei: IFIP SEC 96

Location Management Strategies Increasing Privacy in Mobile Communication

Dogan Kesdogan◊, Hannes Federrath, Anja Jerichow*, Andreas Pfitzmann**

◊Aachen University of Technology, Department of Computer Science Informatik IV, Ahornstr. 55, D-52056 Aachen. E-mail: dogan@i4.informatik.rwth-aachen.de

**Dresden University of Technology, Institute of Theoretical Computer Science, D–01062 Dresden. E-mail: {federrath, jerichow, pfitza}@inf.tu-dresden.de*

Abstract

Mobile communication offers many new opportunities. However, because of the mobility of the subscribers trustworthiness of data, reliability and security are major issues.

Our objective is to increase the network's trustworthiness by providing means to prevent generation of moving tracks: The protection should be shifted into a subscriber‘s domain where the administration of the location information of his mobile station is handled as far as possible. Outside his domain, the subscriber should be able to act anonymously whenever possible.

The location management strategies presented in this paper achieve anonymity of the communicating parties and therefore fulfill the requirement of privacy.

Keywords

Security, Privacy, GSM, Data protection, Anonymity, Pseudonymity

1. INTRODUCTION

Since bandwidth is a very limited resource in radio networks, it is necessary to subdivide the service area into many cells to allow re-use of transmissional frequencies. Therefore, the network structure of GSM [GSM_93] is distributed: several Mobile Switching Centers (MSC) and local databases (Visitor Location Register, VLR) are distributed in the system serving their respective local areas (MSC-Area). The MSC-Area in turn is subdivided into several Location Areas (LA), and an LA is subdivided into several cells.

The smallest unit of the cellular network is the cell. Within a cell the mobile subscriber is able to call anyone and is reachable for anyone. As the subscribers are free to go everywhere in the service area, it is obvious that they will enter and leave cells. Therefore, location information must be managed. Taking some performance considerations into account, the location information maintained in the network is in terms of Location Areas (LA). Currently, management of such data is organized using a central database, the Home Location Register (HLR).

A centralized HLR stores data on subscribers and mobile stations. When a subscriber enters a new LA served by a new MSC, only the relevant data is downloaded to the VLR. For example, the message flow of the Location updating process (LUP) is shown in Figure 1. The mobile station (MS) initiates the LUP (1). The new MSC forwards the request (2) to the HLR which triggers canceling the old record in the database of the previous MSC_{old} (3,4'). In parallel the HLR confirms the updating in MSC_{new} (4) which in turn acknowledges the MS request (5).

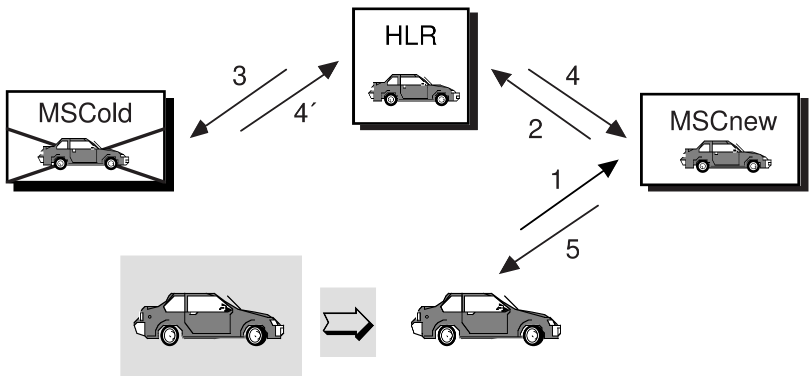

Figure 1 Example of a location update in GSM-network according to [MoPa_92]

As can be seen in Figure 1 the HLR coordinates all data changes. This leads to the security problem of confidentiality. The network provider has unlimited access to the HLR and is able to locate a subscriber using the trace procedure specified in GSM. Hence, at least the network provider is able to (ab)use personal data to record moving tracks. We consider the ability to trace the user of a mobile station and to construct moving tracks without the user‘s knowledge as a violation of the data security requirement "confidentiality".

It is expected, that the number of mobile subscribers will increase dramatically in the future. In order to meet future demands performance of the system must be increased by cell division and reuse of bandwidth. However, with the size of cells becoming smaller and smaller, very detailed information will be available and therefore the determination of the location of a mobile station will become more and more accurate.

Attempts to solve this problem have been proposed in [FJKP_95, Hets_93, Pfit_93, SpTh_93, SpTh_94] and others. We introduce some methods and extend them by new concepts.

The advantages of our new concept include the consideration of privacy, a more efficient administration of data and the usage of the common network structure to a large degree.

2. DATA PROTECTION AND SECURITY REQUIREMENTS

Illegal compilation and unrecognized change of information and disruption of functionality by unauthorized entities in a communication network must be prevented. Related security requirements are availability, integrity and confidentiality.

"Availability" means, that the communication network enables communication between all parties who wish to communicate (and who are allowed to).

"Integrity" relates correctness, completeness and timeliness as well as to the proof that a sender has sent a message and/or the addressee has received it. For the technical fulfillment of integrity, mainly cryptographic methods are applied such as authentication codes and digital signatures.

"Confidentiality" means that data is available to authorized parties only. Especially the protection of the location information of a mobile station needs to be managed. Neither potential communication partners nor third parties (including the network operator(s)) should be able to locate mobile stations or their users. [Pfit_93] suggests the concept of Radio-MIXes[1]. Cryptographic techniques on their own would be insufficient for implementation of confidentiality since our security requirements concern switching data as well (see Figure 2).

security requirement confidentiality	*possible realizations*
message contents	end-to-end encryption
sender and/or recipient anonymity	dummy traffic MIXes
current location of a MS	broadcast
• *location* • *addressing*	new location management strategies
• *protection from locating and identifying a sending MS*	direct sequence spread spectrum (CDMA)

Figure 2 Possible technical realizations of security requirement confidentiality in radio networks

In the past only one aspect has been considered with regard to security. The user of a communication system like GSM has to authenticate himself to the network provider. However, the opposite way has never been an issue.

Yet, we demand *multilateral security*. This means that the security requirements given above have to be guaranteed for everybody no matter whether the violator is a subscriber or a service provider. Confidentiality has to be ensured regarding the individual user, the communication partner and, depending on the application, third parties, particularly service and network providers.

3. TRUSTWORTHY MAINTENANCE OF LOCATION INFORMATION – RELATED WORKS

According to our security requirements the location information of a mobile subscriber must be stored confidentially. One approach is to maintain the location information in a trustworthy environment [Pfit_93, Hets_93]. A personal digital assistant or "personal communication bodyguard" is proposed for managing the location information of a mobile subscriber. The digital assistant in [Pfit_93] is a Home Personal Computer (HPC) located in a trusted and private environment, e.g. the home of the subscriber. The location information of the mobile subscriber is stored in the HPC. If the mobile subscriber roams into another LA the location information in the HPC will be updated.

[1] Radio-MIXes = basic concept for protection of sender, recipient and current location of the subscriber in radio networks which combines the following methods: end-to-end encryption, link encryption between mobile station

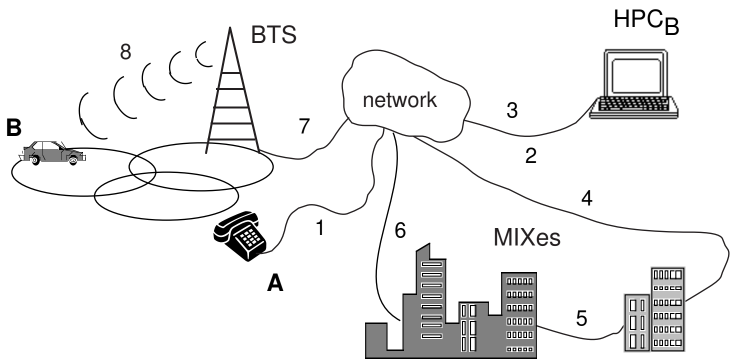

Figure 3 Trustworthy maintenance of location information in a Home Personal Computer (HPC)

In addition, unlinkability between BTSs and HPC (against strong locating attempts) is achieved by MIX-cascades (see [PfPW_91]).

Let us suppose subscriber A wants to set up a call to a mobile subscriber B (see Figure 3). First the network establishes a connection to B's fixed trusted station HPC_B (1,2). The information which BTS is responsible is stored in this HPC_B. Then, the call is routed (through the MIXes‘ unlinkability) to the BTS (3-7). The BTS broadcasts the call in the cell area (8). If B wants to communicate he will establish the connection to A.

In principle the functionality of the HPC (as mentioned above) may be extended to an "intelligent" digital assistant, e.g. a dynamic answering machine and a reachability management machine. In summary that means that each subscriber needs a trusted private device for processing incoming calls (i.e. forwarding, answering etc.).

Every change of the location area leads to LUP-signaling between MS and HPC. If the LA is frequently changed, the signaling overhead for LUPs may increase rapidly. This problem is reduced by doing the following:

At an incoming call the HPC searches the MS by broadcast signaling in the entire network area. Thus, the HPC does not need to manage detailed location information. However, such a system will have a high signaling load. There will also be an optimization problem between location updating and broadcast signaling.

Anonymity by broadcasting makes use of the concept of implicit addresses [PfWa_87]. Implicit addresses describe (in opposite to explicit addresses) neither a location in the network nor a station. Only the intended recipient may recognize that a message is addressed to him. An implicit address is a recognizable mark for the addressee but meaningless and unlinkable for all others.

Visible implicit addresses may be tested for equality by everybody: The subscribers choose arbitrary names (addresses) for themselves, e.g. random numbers. The recipient holds his valid implicit addresses in an associative memory. In this way messages are recognized very efficiently.

Invisible implicit addresses may be tested only by the addressee. For this test a (public key) cryptosystem is necessary. Therefore, invisible implicit addresses are more costly than visible implicit addresses.

If the HPC has the functionality of a reachability manager (see [FJKP_95]) communication demands can be filtered and the overload of the radio network by broadcast be possibly reduced. Another way to reduce the signaling load of a broadcast exists if the HPC is adjustable by the subscriber. In this case the HPC may be adjusted in such a way that the broadcast area is reduced to the area the MS is probably roaming.

It is obvious: the larger a broadcast area the higher the anonymity of everyone.

The planned integration of existing and future cellular systems into UMTS opens up new possibilities for privacy and data protection. The cells (pico-, micro-, macro cells and satellite overlay cells) are at least partly overlaid. Therefore, the term hierarchical cell areas is used.

Suppose there is broadcasting in a possibly large cell area (e.g. satellite cell) even though the MS roams in a small area regarding the subscribed system (e.g. micro cell of GSM). Thus, the necessary location information can be a lot less accurate [FJKP_95]. Therefore, in this approach a wide cell area is used for broadcast depending on the network efficiency and the desired level of anonymity. One technical approach for wide area broadcasting is the usage of low altitude satellites (Low Earth Orbit, LEO).

4. ANONYMOUS SUBSCRIBERS USING PSEUDONYMS

Previous investigations concerning the subscriber's security in location management suggested network structures in which the network operator should not have any information on the subscriber's current location. These approaches imply that the network operator should not have any databases.

In the following chapters we examine the security location management from another angle. Our main concern is that the identity of the mobile user has to be kept secret from the network operator. If we still want to use central databases where the location information is stored then the usage of pseudonyms of the subscriber is a possible approach. Thus, the exact location information is stored but can only be linked to the pseudonym of the subscriber and the network operator can use efficient location management strategies with the aid of central databases.

4.1 Periodic Pseudonyms

As mentioned above the network operator is allowed to keep exact location information if this information is not related to an identifiable subscriber, but to a pseudonym instead. The network is time synchronized and at a time t_i all subscribers in each location area transmit an implicit address to the network.

This implicit address is quite long (e.g. 50 to 100 Bits) and would be treated as a pseudonym by the network administrator. The source of the implicit address is a pseudo random generator (PRG) with the secret key k_B. The same implicit address is generated in the subscriber's MS and in the subscriber's trusted HPC, too. In order to generate the same implicit address in the HPC and the subscriber's MS, there must be a prior key-exchange between these entities. In each time period both MS and HPC create a new pseudonym. The MS in its present location has to inform the network that a new pseudonym has to be registered in this LA. We call this act the pseudonym registration.

Location Registration and Location Updating

The network registers the MS of a subscriber B under a new pseudonym $ID_B = PRG(k_B,t_i)$. The new identity within a location area is called ID_B. It is a newly created entry in the databases of the network. Therefore, the network operator cannot find out the real identity of B although the transmitted information is known, i.e. the ID_B and the location area identifier. $PRG(k_B,t_i)$ is also generated in the HPC of the subscriber. It produces the same ID_B at the same time. At the pseudonym registration the network behaves like a GSM-system when a MS is switched on.

All location updating procedures of the GSM can work under the pseudonym ID_B. Therefore, no modifications in location updating are necessary.

Mobile Terminated Call Setup

With this configuration of the distributed information a call setup leads us to the following procedure (see Figure 4):

Subscriber A wants to communicate to the mobile subscriber B. First, the network needs the current pseudonym ID_B of B. It requests the trusted private environment HPCB of B for ID_B. Against surveillance of B's location the requests have to be audited in the HPC. In the HLR the entry «ID_B,VLR» is used to route the call to the responsible MSC/VLR. The VLR knows the current LA of the pseudonym ID_B and arranges the sending of the paging message.

In relaxing the synchronization requirements, the old IDs in conjunction with the location information should remain longer in the network databases than the time period of the time-synchronization. Since each ID is handled separately, unlinkability of the IDs is guaranteed[2].

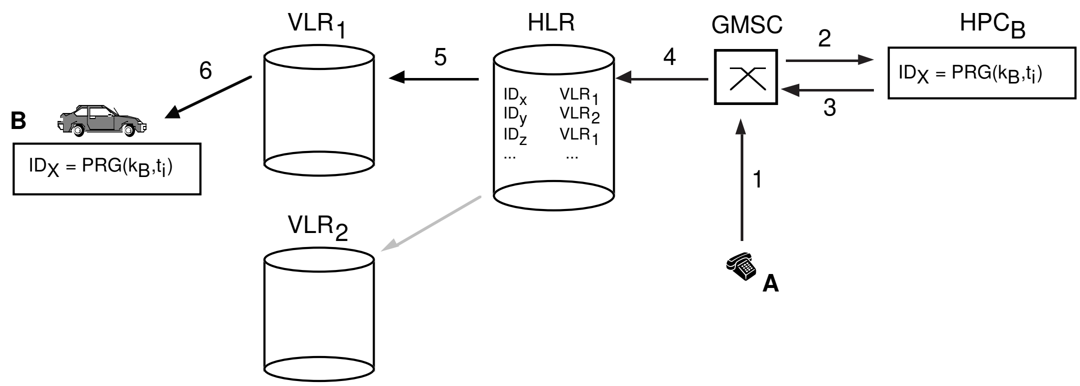

Figure 4 Periodic pseudonyms: Mobile terminated call setup

Compared with GSM we can conclude that the effective additional cost is equivalent to the cost of using a HPC. It can be expressed in terms of more signaling load because we assume that the time period is shorter for the pseudonym method than for the pure location updating. An additional advantage of this is that the system is more robust.

4.2 Group Pseudonyms - An Approach with Location Management according to the logical GSM Structure

The previous approaches required a device in a trusted private environment, e.g. a HPC. The use of a HPC leads us closer to multilateral security, however, it needs technical expense and puts at risk the availability of services (e.g. power failure at home).

If security is organized in centralized environments the mentioned disadvantages of HPCs can be avoided.

The approach suggested in this section requires the following assumption: It must be possible to sample subscribers in subsets (or groups) and to keep anonymous the subscriber in his subset. Then the network operator is allowed to manage the location information of this group. All subscribers of one group are assigned to a group pseudonym.

Now we discuss a technical approach for group pseudonyms: For each subscriber a hash value IMSI' is derived from the IMSI. The hash value IMSI'=h(R,IMSI) (R is a random number) is our group pseudonym for a set of subscribers since various IMSIs are mapped to the same IMSI'.

The logical structure of GSM shall be kept. Therefore, the data bases of our concept are called HLR* and VLR* (according to GSM's HLR and VLR). For illustration see Figure 5.

Location Registration

A mobile subscriber B wants to set up a location registration procedure. By polling the broadcast channel B gets the LA. He sends the hash value $IMSI'_B$ of his $IMSI_B$ to the responsible VLR*.

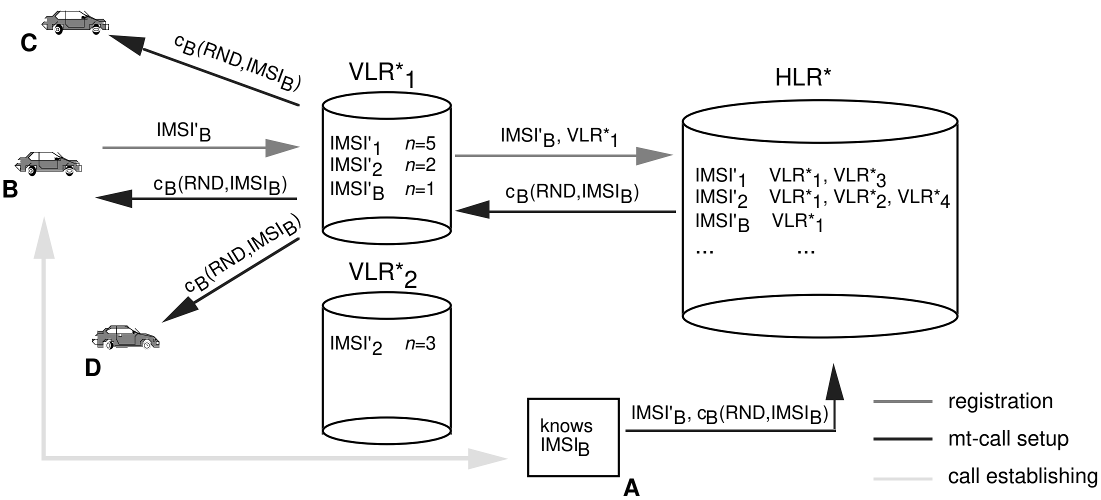

Figure 5 Group pseudonyms: Location registration and mobile terminated call setup

If no subscriber of B's pseudonym group is registered in VLR* a new record «$IMSI'_B$,n» with n:=1 is created. Afterwards, the HLR* is notified that a subscriber of the pseudonym group $IMSI'_B$ is registered.

If an additional subscriber of a registered pseudonym group in VLR* wants to register then the record is updated to «$IMSI'_B$,n:=n+1». In this case HLR* is not notified.

Via n the VLR* can recognize how many subscribers of the same pseudonym group are roaming in the VLR* area.

If the last subscriber of one pseudonym group roams outside the VLR* area, the record «$IMSI'_B$,1» is removed and the HLR* is notified.

In the HLR* the relevant VLR*s are stored to each pseudonym group.

Mobile Terminated Call Setup

If a subscriber A wants to communicate to the mobile subscriber B (mobile terminated call setup, mt-call setup), A calculates $IMSI'_B=h(R,IMSI_B)$[3] and sends $IMSI'_B$ to the network.

Additionally, A encrypts the $IMSI_B$ in the form of an invisible implicit address $c_B(RND,IMSI_B)$. The random number RND guarantees that B cannot be identified by the network. Against attacks by replay of $c_B(RND,IMSI_B)$, a (signed) time stamp sig(T) may be an additional part of the invisible implicit address: $c_B(RND,IMSI_B,sig(T))$.

HLR* sends $c_B(RND,IMSI_B)$ to all registered VLR*s (in our example only VLR^*_1). The VLR*s broadcast the implicit address in the relevant LAs (paging).

By polling the broadcast channel all mobile subscribers receive the messages but only B is able to recognize the message: With his secret key d_B he decrypts the message to $d_B(c_B(RND,IMSI_B))$ and recognizes his $IMSI_B$.

Location Updating

If B roams into another LA a LUP is processed. By polling the broadcast channel he recognizes that a new VLR* is responsible for him. He sends his $IMSI'_B$ to VLR^*_{new}. He arranges the decrement of the counter n (see section Location Registration) in VLR^*_{old} and the increment of n in VLR^*_{new} as well.

Of course, the LUP has to proceed authenticated but this should be already realized by GSM's authentication procedures.

If necessary (see section Location Registration) VLR^*_{old} and VLR^*_{new} notify the changes to HLR*.

---

[3] To enable finding the relevant HLR parts of the IMSI remain unchanged in IMSI', i. e. Mobile Country Code

The presented method keeps the logical structure of GSM to a large degree. However, an essential difference to GSM is the signaling. For the reachability of a mobile subscriber signaling takes place in more than one LA. This aspect increases the average signaling load in the whole network, but not in one LA!

Obviously the critical parameter is the size of the pseudonym group. If the group size is high the danger of tracking of individual subscribers is low, but the signaling load is high and vice versa.

Another difference to GSM is the necessary bandwidth on the broadcast channel for signaling.

By using public key cryptography, one paging message is approximately 500 bit long. However, the extended message space could be used for additional service information (e.g. billing, transmitting of the challenge information for authentication). Possibly, further protocol steps may be reduced.

By using symmetric cryptography, bandwidth on the radio interface is saved, but the key management would be more complicated, since all potential communication partners need to arrange their secret key before they communicate.

5. COMBINING THE ABOVE STRATEGIES

The reviewed and suggested strategies concerning the subscriber's security in location management do not take into account the subscriber's mobility and their geographical distribution. We assume that the centres of population are near to the centres of industry. Most of the time, subscribers are mobile in the vicinity of their homes (usually at home or at work) and are reachable in this limited area. So the number of mobile subscriber decays with the distance to their HPC, as can be seen in Figure 6. New mobile networks will use this distribution of subscribers and provide means to fulfill the need for personal mobility and reachability only within the boundary of these centres. The aim is not the provision of unlimited mobility but serving a large number of subscribers. The effect of this distribution leads to small cell radii and serves subscribers in an economical way. Because of the cost reduction of serving subscribers in geographical limited areas the offered services become cheaper.

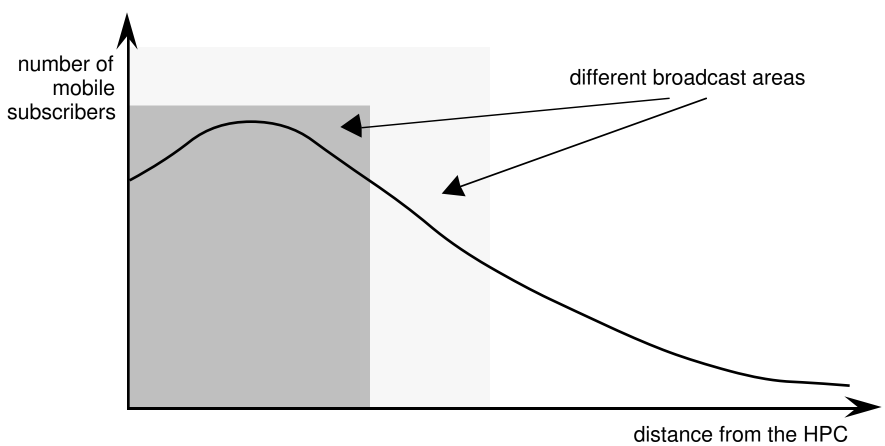

Figure 6 Assumed number of mobile subscribers as a function of their distance from the HPC

In a network with a limited service area it is useful to employ the method of periodic pseudonyms, because the mobile subscriber is in the vicinity of his HPC. The signaling overhead of the request to the HPC is needed in order to find the subscriber's pseudonym. This overhead is economical with respect to the transmission cost in the fixed network because the mobile subscriber is in the vicinity of his HPC.

6. FUTURE WORK

The given location management strategies achieve that even the provider of the mobile communication network cannot trace the roaming of the subscribers. Future work might try to answer the following questions:

(1) How does the technical expense of the location management strategies described in this paper compare with the strategies used so far, which do not aim at protecting the subscribers from observation of their roaming? For different distributions of subscribers and their calls, load -, performance -, and cost models have to be developed and evaluated. As far as possible, a quantification of observation of the roaming should be tried and evaluated in the models as well. For many strategies and some characteristics of roaming and calls we expect that the protection from observation increases the expense.
(2) How could the described location management strategies be improved? These improvements have to be detailed and evaluated as described in (1).
(3) Does it make sense to combine the described strategies, e.g. a strategy of Chapter 3 with a strategy of Chapter 4? These combinations should be evaluated according to (1) as well.
(4) How exactly could and should the paying for services, which are not for free, be organized? Basic protocols for accountability instead of anonymity are known which use digital pseudonyms [BüPf_89, BüPf_90, Pfit_93]. On the one hand, they should be re-thought from the perspective of the new location management strategies. On the other, their expense and security should be determined and compared.

We hope that at least the essential questions can be answered quickly enough to consider the proposed solutions in the further definition of UMTS.

We thank the Gottlieb-Daimler - and Karl-Benz Foundation, Ladenburg, and the German Science Foundation (DFG) for their financial support. For suggestions and discussions, we thank Sven Martin, Jan Müller, Marlies Tischer and Frank Zündorff.

7. REFERENCES

BüPf_89 H. Bürk, A. Pfitzmann: Digital Payment Systems Enabling Security and Unobservability, Computers & Security 8/5 (1989) 399-416.

BüPf_90 H. Bürk, A. Pfitzmann: Value Exchange Systems Enabling Security and Unobservability, Computers & Security 9/8 (1990) 715-721.

FJKP_95 H. Federrath, A. Jerichow, D. Kesdogan, A. Pfitzmann: Security in Public Mobile Communication Networks. Proc. of the IFIP TC 6 International Workshop on Personal Wireless Communications, Verlag der Augustinus Buchhandlung Aachen, 1995, 105-116.

GSM_93 ETSI: GSM Recommendations: GSM 01.02 - 12.21; February 1993, Release 92.

Hets_93 T. Hetschold: Aufbewahrbarkeit von Erreichbarkeits- und Schlüsselinformation im Gewahrsam des Endbenutzers unter Erhaltung der GSM-Funktionalität eines Funknetzes. GMD-Studien Nr. 222, Oktober 1993.

MoPa_92 M. Mouly, M.-B. Pautet: The GSM System for Mobile Communications; A comprehensive overview of the European Digital Cellular Systems; ISBN 2-9507190-0-7; published by Michel Mouly, Marie-Bernadette Pautet.

Pfit_93 A. Pfitzmann: Technischer Datenschutz in öffentlichen Funknetzen; Datenschutz und Datensicherung, DuD 17/8 (1993), 451-463.

PfPW_91 A. Pfitzmann, B. Pfitzmann, M. Waidner: ISDN-MIXes: Untraceable Communication with Very Small Bandwidth Overhead, Information Security, Proc. IFIP/Sec'91, Brighton, UK, 15-17 May 1991, D.T. Lindsay, W.L. Price (eds.), North-Holland, Amsterdam 1991, 245-258.

PfWa_87 A. Pfitzmann, M. Waidner: Networks without User Observability, Computers & Security 6 (1987) 158-166.

SpTh_93 M. Spreitzer, M. Theimer: Scalable, Secure, Mobile Computing with Location Information; Communications of the ACM 36/7 (1993).

SpTh_94 M. Spreitzer, M. Theimer: Architectural Considerations for Scalable, Secure, Mobile Computing with Location Information. Proceedings of the 14th International Conference on Distributed Systems, IEEE 1994.

8. BIOGRAPHIES

DOGAN KESDOGAN received a Dipl.-Inform. degree in computer science from the Aachen University of Technology, Aachen, Germany. He is a research assistant in the mobile communication working group of the computer science departement at the Aachen University of technology. His research activities include third generation cellular systems, mobile and nomadic computing and data security. This work is supported by the Gottlieb-Daimler - and Karl-Benz Foundation.

HANNES FEDERRATH received his diploma degree in computer science from the Technical University of Dresden in 1994. Since then, he works there as a research assistant on security in mobile communications. His main interests are privacy and anonymity in communication networks and cryptography. His work is supported by the Gottlieb Daimler - and Karl Benz Foundation Ladenburg (Germany).

ANJA JERICHOW received her diploma degree in Computer Science from the University of Technology in Dresden, Germany, and her M.Sc. in Computation from Oxford University, UK. Her research interests are in the domain of Theoretical Computer Science. Her studies involved program specification and verification as well as theorem proving. Currently, she is a research assistant at the Department of Theoretical Computer Science at the University of Technology in Dresden. Her work there includes the study of security and privacy, with emphasis on future mobile communication systems. This work is supported by the German Science Foundation (DFG).

ANDREAS PFITZMANN received both his diploma degree in Informatics and his Dr. rer. nat. from the university of Karlsruhe in 1982 and 1989, respectively. Since 1993, he is a professor of computer science at the Technical University of Dresden. His research interests cover privacy and security issues, mainly in communication networks and distributed applications, such as payment systems.