

Distributed Temporary Pseudonyms: A New Approach for Protecting Location Information in Mobile Communication Networks

Dogan Kesdogan^{*}, Peter Reichl, Klaus Junghärtchen
Department of Computer Science • Informatik IV (Communication Systems)
Aachen University of Technology • D-52056 Aachen • Germany
email: {dogan | peter | klausj}@i4.informatik.rwth-aachen.de

Abstract. One of the major security aspects in mobile communication networks concerns information about the localization of the (mobile) network user. This information may be protected by establishing a trusted third party that is responsible for creating suitable pseudonyms for the user identity. Distributing the maintenance of pseudonyms among n independent trusted parties allows increasing further the security of location information. In this paper, a method is proposed that guarantees security as long as at least one of the n parties may definitely be trusted whereas the other parties may turn out to be corrupt. The pseudonym collision probability is derived analytically before a detailed OPNET simulation evaluates the cost of the new approach compared to standard GSM.

1 Introduction

Mobile networks allow users to exchange information or use services independently of their current location. Offering these new possibilities causes a couple of security problems that are significantly more complicated than in the case of fixed networks. Besides the transmission on the air interface, those questions are especially related to the growing intelligence of the networks and the mobility behaviour of the individual users. This paper deals especially with the issue of protecting information about the user location against attacks from inside the mobile network, especially in the case of GSM networks.

GSM divides the supply area hierarchically as follows [19]: The total supply area consists of several MSC areas (each of them managed by a Mobile Switching Center MSC) which are subdivided into several Location Areas LA (each of them managed by a Base Station Controller BSC). Each Location Area finally is constituted of several cells, where a respective Base Transceiver Station BTS emits and receives radio waves within a given frequency spectrum. Within this framework, the location management is distributed: a central database, the Home Location Register HLR contains for each mobile user the respective current MSC area, whereas for each MSC area a local database, the Visited Location Register VLR, contains the current LA in the respective MSC area.

User mobility is managed by the location management functions, especially procedures for mobile originated (MO) calls setup, mobile terminated (MT) calls setup and Location Update (LU). The latter one is an operation performing the update of the user

^{*} The work of Dogan Kesdogan was supported by the Gottlieb Daimler- and Karl Benz-Foundation

location information in the network databases (HLR and VLRs). It is initiated by the Mobile Station (MS) while leaving one LA and entering a new one. If both LAs belong to the same MSC, the procedure is handled locally, otherwise the MSC of the new LA forwards the LU request to the HLR which thereupon cancels the record in the VLR of the old MSC and confirms the insertion of the new record in the VLR of the new MSC which in turn acknowledges the MS request.

If an MT call from the Public Switched Telephone Network PSTN to a GSM subscriber has to be established, the call is routed to the so-called Gateway MSC (GMSC) which requests routing information from the HLR and the respective VLR in order to route the call to the MSC in whose area the MS of the respective user currently is roaming. The MSC then manages the establishment of a radio connection between the MS and the BTS of the cell where the user currently is registered.

There have already been some proposals about how to protect location information in mobile networks (e.g. [8]), especially by using so-called MIXes as introduced by Chaum [2]. Unfortunately, the concept of MIXes has some serious drawbacks, as it implies major changes within the signalling structure as well as the architecture of the mobile network. Moreover, the encryption for the MIX has to exceed 512 bit which adds further load on the air interface. Hence, in this paper a feasible alternative is investigated which has been developed with respect to the following demands:

- protection of the user location information as long as she moves with her mobile switched on but does not receive an incoming call;
- implementation without major changes of the GSM standard;
- sparing use of sensitive data by moving the responsibility for them to the user or to one or several parties the user may trust;
- no restrictions for real-time communication.

To this end, section 2 introduces and discusses the concepts of "Temporary Pseudonyms" and "Distributed Temporary Pseudonyms". The optimal pseudonym length may be derived analytically as demonstrated in section 3. Some major results of a detailed OPNET simulation are presented in section 4, before section 5 concludes the paper with some final remarks. Note that the companion paper [20] is focussed towards the performance evaluation aspects of the proposed method and contains a lot of simulation results in far more detail.

2 The Method of Distributed Temporary Pseudonyms

The basic idea of "Temporary Pseudonyms" (TP) has been proposed first in [14, 16]. After introducing, refining and discussing this concept in section 2.1, the following section 2.2 presents an important extension that allows to distribute the pseudonym information among several trusted parties without significantly increasing the signalling load. These concepts have been implemented and evaluated by extensive simulations which finally will be demonstrated in section 4.

2.1 The TP Method

Temporary Pseudonyms. The method of Temporary Pseudonyms (TP method) is based on the concept of trusted parties where e.g. a Home Personal Computer (HPC) *confidentially* stores sensitive data (authentication keys, location information etc.) or even handles the complete location management (thus replacing the VLR) [7]. But in

our context the trusted region is no longer used for saving the actual location information of the user. Instead, the basic idea of the TP method consists of protecting a mobile user's location information (within a mobile network) by protecting her identity. To this end, the user is assigned a pseudonym (Pseudo Mobile Subscriber Identity PMSI). As long as the user is registered under a pseudonym, the network provider may know that a user under a certain pseudonym currently is at a certain place, but he is not able to link the user's real identity with her present location.

There are different possibilities for choosing a suitable trusted region. In the case of the TP method, the trusted region consists of a HTD (Home Trusted Device) that is completely under the control of the user. The pseudonym is generated in the MS and the HTD in a time-synchronous manner. To this end, in both stations two identically parametrized pseudo-random generators PRG generate a random number that is between 50 and 100 bits long. The seed for initializing the PRG has previously been exchanged between MS and HTD. The resulting random number is used as an *implicit address*, i.e. works as the "identity" of the user, under which she registers in the data bases (HLR and VLR) of the mobile network. Within the procedures in the GSM network, the implicit address plays the role of the IMSI (for a detailed discussion of implicit addressing see [6, 13]).

If a mobile terminated call is arriving at the network, the HTD of the user – which may be reached via the ISDN fixed network – is asked for the currently used pseudonym. With this information the network provider is able to find out the location information and to set up the call as usual. Note that the user remains anonymous as long as no call arrives for him from outside. Moreover, mobile originated calls may happen completely anonymously as the user location need not to be determined, if the user uses prepaid phone cards. In order to guarantee the encryption of the data sent via the air interface during an MO call it turns out to be necessary that the MS itself informs the network provider about the (symmetric) key that is used for communication (as an unambiguous secret key would allow to deanonymize the user). This last step may take place using an asymmetric encryption scheme (RSA). The user encrypts the key to be used with the public key of the network provider.

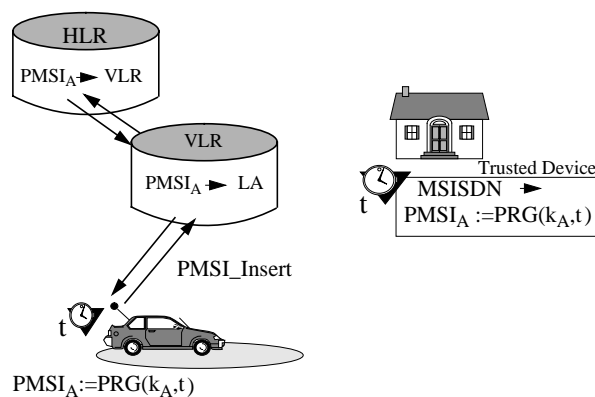


Fig. 2.1. Generation and Insertion of a New Pseudonym

Changing the pseudonym takes place time-synchronously between MS and HTD (cf. fig. 2.1). The most efficient way might be changing the pseudonym exactly while

changing the location area, i.e. by registering in the new location area under a different pseudonym (without cancelling the old one, because otherwise the two pseudonyms could be linked). This idea hides any information about a direction of the pseudonym, but on the other hand the user may step out of her anonymity (as e.g. a constant driving speed yields constant pseudonym changes and thus linkability). Moreover, it is not possible to directly inform the HTD about the pseudonym update as every message between the user and her HTD lets the user's identity be combined with her location area. Therefore, the time between two pseudonym changes is drawn from an exponential distribution and thus memoryless. Additionally, this approach prevents situations in which all users have to update their pseudonyms at the same time, which easily could exceed the channel capacity of the mobile network.

Discussion of the TP Method.

- *Passive Attacks.* The security of the TP method is based on the pseudonym and the HTD of the user. As long as the user has her MS switched on but does not communicate, this method hides her location information against the passive attack of an insider, i.e. the attempt to get location information about the user by observing the network data bases (HLR and VLR). As soon as the user receives a call, her location (i.e. the LAI) is revealed as her HTD reveals her current pseudonym to the network provider who then is able to link user identity and location. Hence, this passive attack cannot be prevented by the TP method. If the insider observes these location information over long time, he might be able to generate a movement profile for the user.
- *Active Attacks.* Active attacks of the network provider, i.e. attempts to find out the user location by periodically asking her HTD, may be recognized because all demands are logged at the HTD. Hence, if there are much more demands at the HTD than actual calls, this points towards an active attack. Note that the user may recognize this attack but is not able to prevent it. One solution might be to add the functionality of a reachability manager to the HTD, i.e. it could decide for each demand whether the importance of the demand justifies revealing the pseudonym.
- *Real-time Communication.* The most important advantage of the TP method compared to other methods for securing the location information consists of its easy and efficient implementation within existing mobile networks and the small additional signaling overhead. Moreover, the possibility of real-time communication is not restricted.
- *Availability.* The major availability problem consists of the possibility of a HTD breakdown. In this case, its availability could only be restored by the user herself which might be difficult in case of her absence. Moreover, if the user is not informed separately about the breakdown, she will not notice it until another user who wanted to call her sends her an extra notification (or she tries to call herself).

TP Method with a Trusted Party. A rather straightforward extension of the TP concept prevents the availability problem by putting the pseudonym generation into the hands of a third party (e.g. an outside company) which the user may trust. In this case, it is possible to exchange direct messages between MS and the trusted device (TD). As the TD administrates a lot of users, a message from the MS may reveal the relation between the pseudonym (MS) and the group of users administrated by the TD, but not the identity of the user herself. Hence, changes of the pseudonym change rates

may be sent directly to the TD, moreover the availability of the TD may be increased compared to the HTD approach. The remarks about active and passive attacks remain the same as in the case of HTD. Note especially that a collaboration of the (unconditionally trusted) TD with the network provider is sufficient for revealing the location information to the insider.

2.2 The DTP Method

In contrast to the TP method, DTP distributes the generation of the pseudonym towards n different trusted parties. Each party i holds for each user the following entries:

- *MSISDN*, the phone number of the mobile user;
- *pseudo*, the seed for the pseudo-random generator of the pseudonyms;
- Δt , the time until the next pseudonym change takes place;
- $pmsi_i$, the part of the pseudonym *PSMI* generated by party i .

At certain points in time, each trusted party generates its new partial pseudonym $pmsi_i$ ($1 \leq i \leq n$), where no party knows any of the pseudonyms of the other parties. At the same time, the MS registers at the data bases of the mobile network under the (total) pseudonym

$$PSMI := pmsi_1 \oplus pmsi_2 \oplus \dots \oplus pmsi_n, \quad (2.1)$$

where " \oplus " corresponds to the XOR function.

As soon as a call arrives at the gateway (GMSC), the n trusted parties are asked for their "partial pseudonyms". The answers then are XORed bit by bit yielding the actual pseudonym *PSMI* of the user as shown in fig. 2.2.

The idea of a bitwise superposition of messages follows the concept of DC networks as introduced in [3], but instead of guaranteeing the sender anonymity, the aim of our approach is to distribute the information about the user pseudonym over n independent parties such that the information is safe as long as at least one of them works indeed correctly instead of collaborating with the network provider.

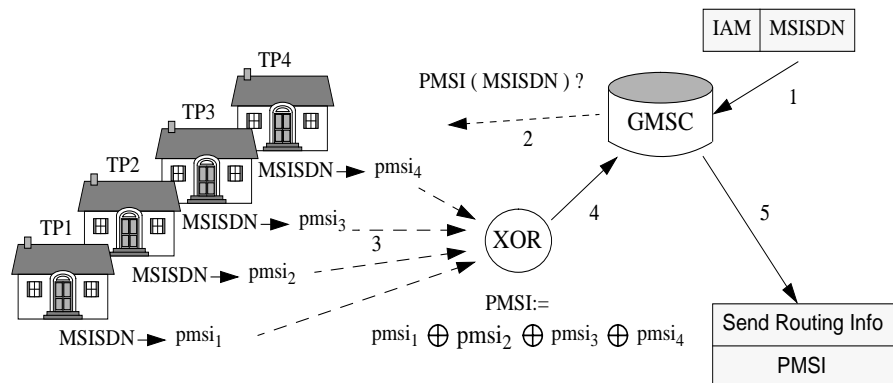


Fig. 2.2. Signalling of Distributed Temporary Pseudonyms

Finally the user may be found as usual over her pseudonym, and the call may be set up. Note that as in the case of the TP method, the user is deanonymized as soon as a mobile terminated call arrives, and her location may be found by the insider. If the time for the next pseudonym change would be determined unambiguously for each user and told to each of the trusted parties, a corrupt "trusted" party might be able to inform the insider beforehand about the moments of changes. In this case, the insider only has to observe in which location area there is a pseudonym change at that time in order to know the relation between identity and location. Hence, it is necessary to either hide the time of the next pseudonym change within a number of synchronous changes of partial pseudonyms or to weaken the knowledge of a trusted party i about the fact that changing its partial pseudonym $pmsi_i$ indeed yields a change of the total pseudonym $PMSI$ in all of the data bases. These deliberations yield the concepts of "change rate classes" and "synchronized pseudonyms", resp.

Change Rate Classes. In order to maintain large sizes of the anonymity sets it is optimal if all users change at the same time, as in this case the information about the time of the next pseudonym change is of no value at all. But this obviously is impossible because of the limited channel capacity.

Hence, the networks proposes j possible "change rate classes" (e.g. approx. 20 sec, 1 min, 20 min etc.). The j classes should be chosen such that the users of a network are distributed as uniformly as possible over the classes. All members of a certain class change their pseudonyms at the same time. The MSs determine themselves (e.g. using the approximate rate at which they change location areas) the class they wish to belong to. Thus in the case of deanonymization of a user, little is revealed about her prior movement path. The less classes there are and the more members one class possesses, the larger is the "anonymity set" (cf. [15]) of the resp. user (but the larger is the signalling amount at change time, too). An entry change (e.g. the change to a new class) at a trusted party then is performed e.g. while the user is deanonymized.

Synchronized Pseudonyms. Here, the trusted parties are subdivided into two parts. Each party is synchronized with exactly one other party (without knowing it), i.e. their initialization of the PRG and the pseudonym change times are identical (w.l.o.g. assume that for each pair, one party is between number 1 and $n/2$, whereas the companion is between $n/2+1$ and n). The only one to know which parties are synchronized and which initializations the pseudo-random generators of the single parties possess is the user.

The idea is to prevent a party from being able to predict that the actualization of its partial pseudonym yields indeed a change of the total pseudonym. This is reached because the partial pseudonyms of two parties may cancel each other out because of the XOR function. Each trusted party i possesses the following entries per user:

- $MSISDN$, the phone number of the mobile user;
- $pseudo_i$, the seed for the PRG of the pseudonyms;
- $update_i$, the seed for the PRG of update times;
- $uniform_i$, the seed for the PRG of uniformly distributed random numbers;
- p_i , the probability that a generated partial pseudonym will indeed be used;
- uni_i , a value randomly drawn from a uniform distribution;
- Δt_i , the time until the next pseudonym change takes place at party i ;
- $pmsi_i$, the part of the pseudonym $PMSI$ generated by party i .

The concept of synchronized pseudonyms yields that exactly two parties (i and $n/2+i$, say) have identical entries. At certain moments, each trusted party generates its new partial pseudonym $pmsi_i$ (no party knowing the pseudonym of any other one nor the moment any other party generates a new partial pseudonym). Additionally, each party determines whether $pmsi_i$ will be used by comparing the random number uni_i with the limit p_i : If the random number drawn from a $[0;1[$ -uniform distribution is smaller than p_i the new partial pseudonym is used (for the sake of simplicity let us assume that all $p_i = p$ are identical). Note that each time one pseudonym is generated, a different trusted party generates the same partial pseudonym with probability p . As the total pseudonym is only changed if exactly one of the two new partial pseudonyms is used, with probability $p^2 + (1-p)^2$ the two parties behave identically (i.e. the newly generated pseudonym has no effect), and with probability $2 \cdot p \cdot (1-p)$ the user registers under a new total pseudonym.

Now, the expected time Δt between two changes of the total pseudonym can be calculated as

$$\Delta t := \min\{\Delta t_1, \Delta t_2, \dots, \Delta t_{n/2}\} \quad (2.2)$$

with

$$\begin{aligned} \Delta t_i &:= PRG_{update_i}, \quad 1 \leq i \leq n \quad \text{and} \\ \Delta t_i &:= \infty \quad \text{if } (uni_i < p \wedge uni_{i+n/2} < p) \quad \text{or} \quad (uni_i \geq p \wedge uni_{i+n/2} \geq p) \end{aligned} \quad (2.3)$$

where (2.3) describes the case that the total pseudonym does not change.

Assuming that each trusted party ($1 \leq i \leq n$) draws Δt_i from an exponential distribution with parameter λ , a straightforward calculation shows that for identical acceptance probabilities $p_i = p$, Δt can be characterized as drawn from an exponential distribution with parameter

$$\sum_{i=1}^{n/2} (2 \cdot p \cdot (1-p)) \cdot \lambda = p \cdot (1-p) \cdot n \cdot \lambda. \quad (2.4)$$

Hence, the time until the next change of the total pseudonym is still memoryless, i.e. an attacker is not able to predict it. Choosing p very high (or low) implies that each trusted party has only small knowledge about the next total pseudonym-update which makes it even harder for the attacker to explore it.

Finally it should be noted that the idea of synchronized pseudonyms requires a change in the attacker model, as at least one synchronized pair out of the n parties is required to work correctly.

3 Analysis of Pseudonym Collision Probability

3.1 Length of the Pseudonym

Within the TP method, the pseudonyms consist of random numbers (generated in the HTD and the MS) which should be different for all users. In order to hide as much user movement information as possible from the attacker, the pseudonym changes take place according to a Poisson stream whose parameter corresponds to the location area change rate. In analogy to [9], assume A pseudonyms are already registered in the HLR of the network operator (with $A \geq N$, the number of users). To calculate the

average time T_{coll} until two pseudonyms coincide if this procedure is repeated independently, let λ_n be the rate of pseudonym changes (identical for all users). Hence the pseudonym changes of all N users form a Poisson stream of rate $\lambda = N \cdot \lambda_n$ with interarrival time $1/\lambda$.

If the pseudonyms are B bits long, the probability of choosing an already existing *PMSI* is $A/2^B$, hence the random variable X , which describes the number of pseudonym changes until the first collision, is distributed geometrically with distribution function $P\{X = k\} = \frac{A}{2^B} \cdot \left(1 - \frac{A}{2^B}\right)^{k-1}$ and expectation value $E(X) = \frac{2^B}{A}$.

Therefore, the mean time between two collisions may be calculated as

$$T_{coll} = \frac{E(X)}{\lambda} = \frac{2^B}{A \cdot N \cdot \lambda_n}. \quad (3.1)$$

According to (3.1) it turns out that (in the case of 10 million active users) for pseudonym lengths of 100 bits the expected time between two collisions is around 20 billion years which seems to be sufficient for our purposes. Note on the other hand that a length of 50 bits yields a time of only 7 days whereas 70 bits yield around 16 years. Therefore, in the OPNET simulation as described in section 4, a pseudonym length of 100 bits has been used throughout.

3.2 Average Number of Pseudonyms in the HLR

Assume the interarrival times of pseudonym changes of each user to be exponentially distributed with mean $1/\lambda_n$. Pseudonyms are created independently of each other, hence the total rate of pseudonym changes in the network is $\lambda = N \cdot \lambda_n$ and the interarrival time between each two changes is exponentially distributed with rate $1/\lambda$.

In order to prevent the HLR to be overloaded by pseudonym entries, each pseudonym is assigned a "time to live" parameter TTL. Thus, the total time a pseudonym entry is stored corresponds to the sum of the time until the next pseudonym change and the TTL. In order to prevent an attacker from drawing any conclusions from the moment of a pseudonym deletion and the moment of storing a pseudonym (which might allow him to link both events and hence to link both pseudonyms), the TTL parameters are drawn from an exponential distribution with parameter μ_{TTL} . Therefore, the probability of deleting a pseudonym is independent from the moment of registering it due to the memoryless property of the exponential distribution.

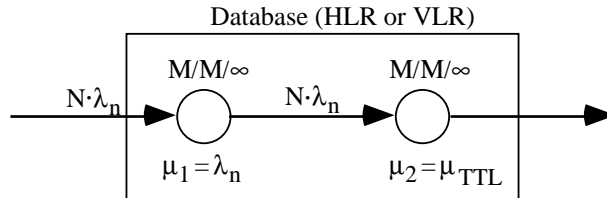


Fig. 3.1. Tandem Network

The HLR may be modelled as "tandem network" (fig. 3.1) consisting of two subsequent $M/M/\infty$ -queueing systems. The assumption of infinitely many servers corresponds to all pseudonyms held independently from one another in the HLR until their deletion. The first server delays all entries until a new pseudonym is generated, i.e. for each user with service rate $\mu_1 = \lambda_n$. The second server delays the entries by the time to live, hence its rate is $\mu_2 = \mu_{TTL}$. As the queueing systems of a tandem network with these properties may be viewed independently from one another, the arrival rate at the two queueing systems of the tandem network equals $N \cdot \lambda_n$.

The average number of customers in an $M/M/\infty$ -system equals $\rho = \lambda/\mu$ [17]. Hence, the average number of pseudonyms in the HLR may be calculated as

$$A = \rho_{total} = \rho_1 + \rho_2 = \frac{N \cdot \lambda_n}{\lambda_n} + \frac{N \cdot \lambda_n}{\mu_{TTL}} = N \cdot \left(1 + \frac{\lambda_n}{\mu_{TTL}} \right). \quad (3.2)$$

If the TTL of a pseudonym is drawn from an exponential distribution with the same parameter as the interarrival times of pseudonym changes, the average number of HLR entries equals $A = \rho_{total} = 2 \cdot N$, i.e. twice the number of users.

4 Simulation Results

4.1 Preliminaries

The TP method has been integrated in microscopic detail into an existing GSM model for the simulation platform OPNET [5]. The model consists of the OPNET network layer for building the network topology, the OPNET node layer for modelling the data flow within network components, and the OPNET process layer, in which the behaviour of the components is characterized by finite automatas. In this model, the main GSM hardware and signalling procedures are simulated (see fig. 4.1). Included are mobile stations (MS), base transceiver stations (BTS), base station controllers (BSC), mobile switching centers (MSC), visited location registers (VLR), home location registers (HLR), gateway mobile switching centers (GMSC) and the public switched telephone network (PSTN).

Signalling has been modelled close to the GSM standard down to layer 3 of the protocol model. In addition to [5], packets sent over the SS7 network are piggybacked by appropriate packets of the lower layers TCAP, SCCP and MTP3.

The simulation comprises the procedures Mobile Originated Call, Mobile Terminated Call, (Inter-BSC, Inter-MSC) Location Update, (Inter-BSC, Inter-MSC) Handover and Paging. Service handling and authentication are not integrated, in contrast to System Information and Measurement messages, which are evaluated for deciding on the initiation of location updates and handovers, resp.

The simulation area consists of 12 cells, 4 location areas, 2 MSC/VLRs and 21 mobile users. Each location area consists of 3 cells. Larger networks with up to 40 cells, 4 MSCs and 50 mobile users have been tested, the restriction to the smaller one is due to simulation duration and consistency of the results. The diameter of the cells depends on the mobility scenario and has been chosen to be 1.5 km for a city scenario, 15 km for countryside, 10 km for "Autobahn" (i.e. German Highway) and 9 km for a mixed scenario.

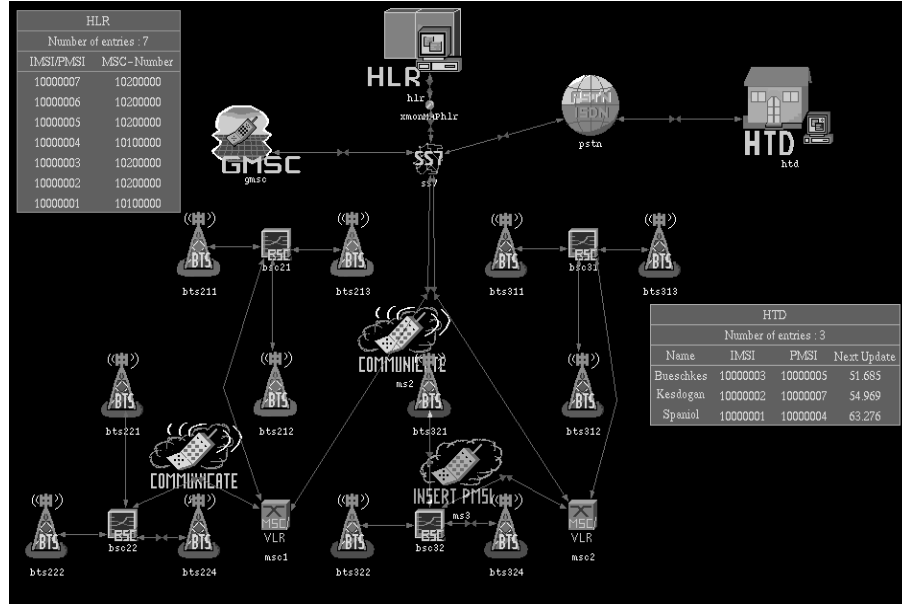


Fig. 4.1. The OPNET model for the simulated GSM network

The mobility model for the MS has been implemented according to results from [4], who measured delays in bus traffic and found out that these delays could be modeled by an Erlang-4-distribution. The bus delays mainly occur due to the individual traffic sharing the roads with the buses. Though based on slightly different driving conditions (e.g. time schedule at every bus stop), the results can be transferred to individual traffic. Hence, the mean delay may be calculated as "perturbation" of the usual mean driving time according to

$$t_{mean} = t_{min} + X_{Erl} \cdot t_{min} = (1 + X_{Erl}) \cdot t_{min} \quad (4.1)$$

where X_{Erl} is drawn from an Erlang-4 distribution. This yields a mean simulated speed of

$$v_{mean} = \frac{v_{max}}{1 + X_{Erl}}. \quad (4.2)$$

This leads to the following mobility algorithm that is executed for every mobile node during the simulation:

1. Mobile users are distributed randomly across the simulation area.
2. Each mobile user draws a random destination from a uniform distribution.
3. Each mobile node has parameters indicating the real mean speed and the maximum speed. Thus the delay for part of the way and the simulated speed may be calculated.
4. At certain times each mobile node decides whether it performs a 90° turn towards the destination or not. The distance between these points and the turn probability are given as parameters and depend on the simulation scenario; they have been calculated based on typical respective scenarios in the region of Aachen.
5. The mobile node moves according to its calculated speed and direction.

6. If the destination is reached, a new destination is chosen according to step 2, and the algorithm continues from step 3 onwards.

Note that the mean speed of mobile nodes has been assumed to be 10 km/h in the city, 65 km/h in the countryside, 100 km/h in the "Autobahn" and 58 km/h in the mixed scenario (cf. [18]).

Measurements in real mobile networks [1] state that during the busy hour approximately $2/3$ of the calls are mobile originated whereas $1/3$ are mobile terminated calls. Measured mean call setup rates are around 0.45 calls/hour for mobile originated calls and 0.225 calls/hour for mobile terminated ones, yielding a global call interarrival time of about 5333 seconds. Furthermore, the mean duration of mobile terminated calls has been measured to be 115 seconds in contrast to 105 seconds for outgoing calls. According to [10, 11], call holding times as well as interarrival times may be drawn from exponential distributions.

4.2 Variation of the Mobility Behaviour

In this section it is demonstrated how changes in the mobility behaviour effect the load of the mobile network. To this end, the three scenarios "city", "countryside" and "Autobahn" are explored, using typical simulation parameters in each case.

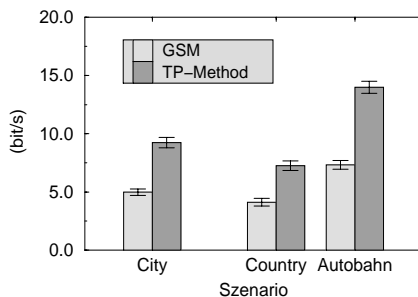


Fig. 4.2. Total Network Load in the three Scenarios

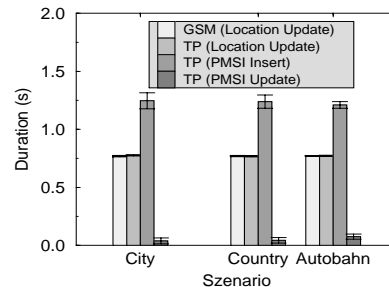


Fig. 4.3. Mean Duration of the Procedures for Location Update and Pseudonym Change

Fig. 4.2 shows the total signalling load generated by one user in the mobile network. On the one hand, the network load increases proportionally to the mobility environment of the user. A user in the countryside scenario with large location areas and medium speed causes less load than a user who is fast or moves in areas with small cell sizes. This demonstrates the influence of the rate of location area updates. On the other hand, the difference between the signalization in the GSM network with and without the TP method appears to increase proportionally to the mobility behaviour as well, being 75% in the countryside, 80% in the city and 96% in the Autobahn scenario.

The mean duration of the procedures for location update and pseudonym change is demonstrated in fig. 4.3. On the air interface the location update procedure uses the

TMSI for either method. Thus the length of the PMSI here is of no influence at all. Within the SS7 network PCM30-connections are used. The 64 kbit/s speed of their signalling channel is high enough to prevent the location update delay from rising heavily. The PMSI Insertion procedure has to send one pseudonym across the air interface. Due to its length and the low rate of the dedicated signalling channel the duration of the procedure increases by 67%. Nevertheless it is not as time-critical as a handover procedure.

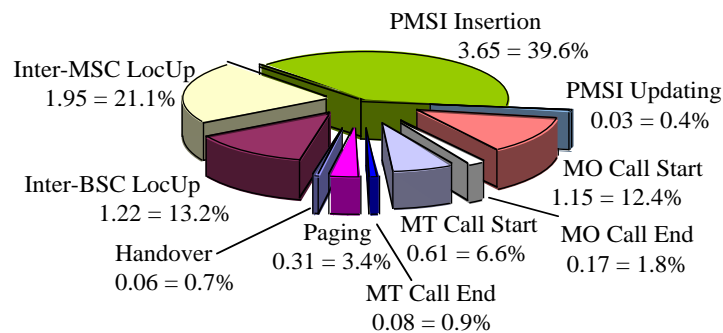


Fig. 4.4. Distribution of the global load w.r.t. the different GSM procedures (in bit/s) for the case of the city scenario

Where does the high signalling load come from? The responsible factor turns out to be the share of the location update procedure, which is extended for the PMSI insertion procedure, in the global load. Fig. 4.4 shows that in the chosen simulation environment and with the TP method being implemented, location updates and pseudonym insertions amount to 74% of the global load, whereas the low impact of the call procedures results from the call rates being adjusted to the values given in [1].

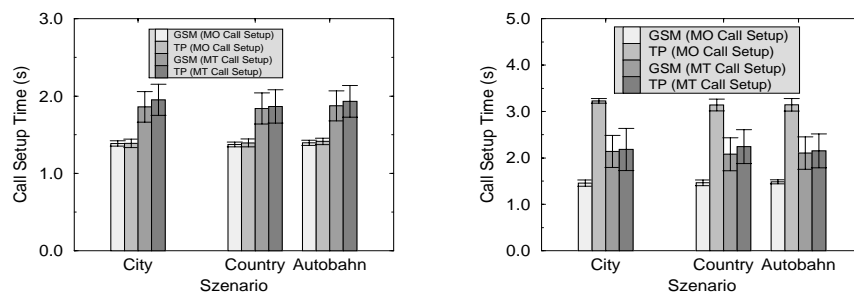


Fig. 4.5. Duration of the Call Setup without (left) and with (right) encrypting a session key for Mobile Originated calls

Fig. 4.5 finally shows the call setup time for incoming and outgoing calls. The call setup times for both kinds of calls coincide as long as no (public key encrypted) session key is sent from the mobile to the network. Conveying a session key causes a delay of 0.66 seconds per direction if sent over a dedicated signalling channel with netto bit rate of 782 bit/s. The pseudonym has to be sent twice with an initial mes-

sage in order to prevent two users from accessing the same traffic channel: First a random number is sent from the user to the network which afterwards is contained in the answer of the network to indicate that the user is now owner of the channel which is assigned by the network. Then the initial message is sent "piggybacked" to a layer 2 LAPDm SABM-message which is answered by an exact copy of it within an LAPDm UA-message (see [19]). If the mobile does not receive an exact copy of its initial message, it has to clear the connection.

4.3 Variation of Pseudonym Change Rate

This section explores the influence of the pseudonym change rate on the load of the GSM network. We have investigated pseudonym change intervals from two minutes up to one hour. The mobility and call model is the same as in the previous section. Fig. 4.6 left shows that the shorter the interval between two pseudonym changes, the larger the increase in the global load and vice versa. Note that the load may be influenced to a much larger extent than e.g. by the call rate (or PMSI length, as further experiments have shown). Fig. 4.6 right demonstrates how the PMSI insertion procedure becomes more and more responsible for the total network load: At rate 600 changes/hour the insertion causes 47% whereas at rate 120 changes/hour this share is already approx. 81%. The reason for this lies in the "length" of the PMSI insertion procedure (i.e. the sum of lengths of all respective messages) which is many times the length of the PMSI. Hence it can be concluded that increasing the PMSI exchange rate has much more effect than prolonging the PMSI.

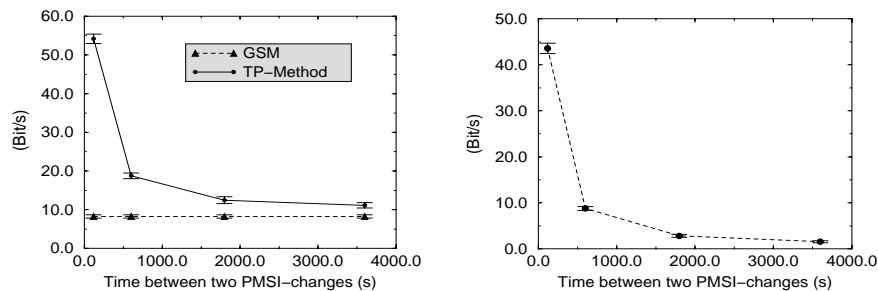


Fig. 4.6. Global Load for Different PMSI Change Rates (left) and Load Caused by the PMSI Insertion Procedure (right)

Fig. 4.7 shows the distribution of the total load w.r.t. some interfaces. Obviously, the air interface is affected most by the signalling of a user. As the dedicated signalling channels at the air interface are assigned exclusively to the users, whereas the signalling messages of all users use the resources of the same SS7 network, the load in the SS7 network must be viewed in relation to the user number. A load increase of 0.1 bit/s (as caused by increasing the pseudonym change rate from 2 to 6 changes/hour) increases the total load in the SS7 network for a user number of 1 million by approximately 98 Kbit/s.

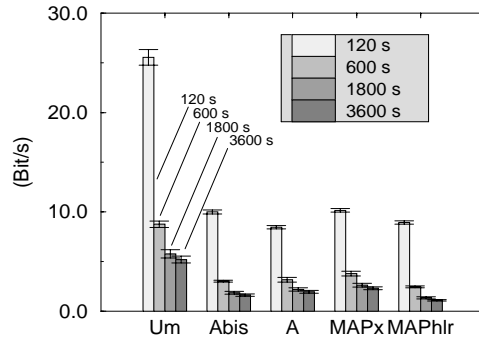


Fig. 4.7. Load at Different Air Interfaces for Different Pseudonym Change Rates

There have been conducted a lot more OPNET simulations, e.g. investigating the influence of varying the pseudonym length or call rate. The results of these simulations are reported in great detail in [12, 20].

4.4 Application of the Results

From the results demonstrated in the previous sections it may be concluded that especially the pseudonym changes have major influence on the network load. The TP method as investigated so far is fixed to the location update procedure in order to guarantee the reachability of the user under all circumstances. It is complicated to perform the change of pseudonyms exactly at the boundary of the location area because the user's HTD does not easily know the exact moment. But it seems worth considering to perform the change of location areas in direct conjunction with the change of pseudonyms as – compared to the TP method – this might save the expense of the location update procedure. In this case it must be noted at which rate incoming call attempts will fail as the user cannot be reached.

A second point for improvements is the sending of the transmission key in the case of outgoing calls. As explained above, an exact copy of the initial call setup message is sent back to the user in order to prevent several users from using the same dedicated channel. This sending doubles the cost for introducing the transmission key.

The following simulation investigates the effect of the above improvements. Fig. 4.8 demonstrates that the total load may be reduced by 17% compared to the TP method, whereas the load at the air interface decreases by 22%. This may be explained by the fact that the 512 bit transmission key is only sent once per initial call setup message. As the pseudonym change is performed exactly as with the original TP Method, the improvement at the HLR interfaces comes from avoiding the inter-MSC-location-update procedure. Sending the transmission key only once reduces the time for setting up a call by approx. one second (fig. 4.8 bottom right).

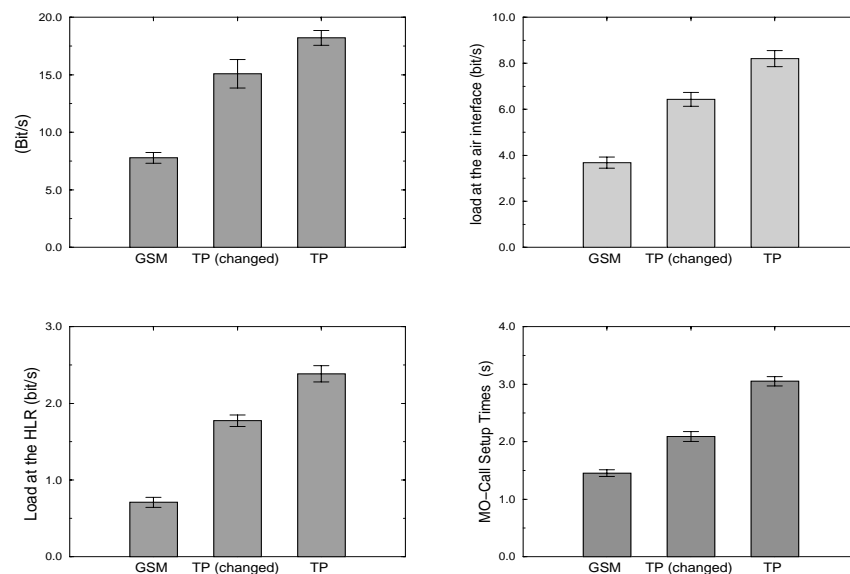


Fig. 4.8. Global load, load on the air interface (top row) and the HLR Interface and call setup times (bottom row) for the different methods

Note that the enhanced TP method does not guarantee the strict reachability of the user as the location updates are replaced by pseudonym changes: if the time between two pseudonym changes becomes too long, possible location area changes are not communicated to the network data bases. Therefore, in our simulation about 17% of call attempts were not able to be performed successfully. This value decreases with the user mobility. As a consequence, either the pseudonym change rate must increase (according to the mobility behaviour of the users) or another paging strategy must be introduced, e.g. the method of dynamical location areas.

5 Concluding Remarks

This paper has dealt with the question of how to protect information about the user location against insider attackers in a GSM network. It has been demonstrated that the TP and DTP methods yield feasible solutions to this problem by hiding the relation between user identity and location using pseudonyms whose unavoidable changes can be protected from being linked together by an attacker. However, even with this approach the mobile user will be deanonymized as soon as she accepts any incoming call. Current and future research deals with this issue, especially with the question whether it is possible at all to conduct completely anonymous calls.

References

- [1] Brass, V.; Fuhrmann, W. F.: Traffic Engineering Experience from Operating Cellular Networks. *IEEE Communication Magazine*, August 1997, 66–71.
- [2] Chaum, D.: Untraceable Electronic Mail, Return Addresses and Digital Pseudonyms. *Communications of the ACM* (24) 2, 1981, 84–88.
- [3] Chaum, D.: The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability. *Journal of Cryptology* vol.1 no. 1, 1988, 65–75.
- [4] Dauber, F.-J.: QoS Parameters of Bus Traffic with respect to Time-Tables and Subsequent Variability. Ph.D. Thesis (in German). Aachen University of Technology, July 1986.
- [5] Institute Eurecom, Delta Partners: A GSM OPNET Model (incl. documentation). © 1994–95.
- [6] Farber, D. J.; Larson, K. C.: Network Security via Dynamic Process Renaming. *Proceedings of the 4th Data Communications Symposium, Quebec (Canada)*, Oct. 1975, 8/13–8/18.
- [7] Federrath, H.; Jerichow, A.; Kesdogan, D.; Pfitzmann, A.: Security in Public Mobile Communication Networks. *Proc. of the IFIP TC 6 International Workshop on Personal Wireless Communications. Augustinus (Aachen) 1995*, 105–116.
- [8] Federrath, H.; Jerichow, A.; Pfitzmann, A.: MIXes in Mobile Communication Systems: Location Management with Privacy. *Proceedings of the Workshop on Information Hiding, Cambridge (UK)*, May 1997.
- [9] Federrath, H.: Trusted Mobility Management in Telecommunication Networks. Ph.D. Thesis. TU Dresden 1997 (in German).
- [10] Fuhrmann, W.; Brass, V.: Performance Aspects of the GSM Radio Subsystem. *Proceedings of the IEEE*, vol. 82 no. 9, 1994, 1449–1466.
- [11] Guérin, Roch A.: Channel Occupancy Time Distribution in a Cellular Radio System. *IEEE Transactions on Vehicular Technology*, vol. VT-35 no. 3, 1987, 89–99.
- [12] Junghärtchen, K.: Simulative Investigation of Methods for Protecting Location Information in Mobile Networks. Diploma Thesis. Aachen University of Technology, November 1997 (in German).
- [13] Karger, P. A.: Non-Discretionary Access Control for Decentralized Computing Systems. M.Sc. Thesis, Techn. Report MIT/LCS/TR-179, MIT 1975.
- [14] Kesdogan, D.; Foulletier, X.: Secure Location Information Management in Cellular Radio Systems. *Proceedings of the IEEE Wireless Communication Systems Symposium WCSS'95, Long Island, 1995*, 35–46.
- [15] Kesdogan, D.; Egner, J.; Büschkes, R.: Stop-And-Go-MIXes Providing Probabilistic Anonymity in an Open System. *Workshop on Information Hiding. Oregon, April 1998* (to be published in Springer LNCS).
- [16] Kesdogan, D.; Federrath, H.; Jerichow, A.; Pfitzmann, A.: Location Management Strategies increasing Privacy in Mobile Communication Systems. *Proceedings of the 12th IFIP International Information Security Conference SEC'96, May 1996* (Chapman & Hall).
- [17] King, P. J. B.: *Computer and Communication Systems Performance Modelling*. Prentice Hall 1990.
- [18] Lyberopoulos, G. L.; Markoulidakis, J. G.; Polymeros, D. F. et al.: Intelligent Paging Strategies for Third Generation Mobile Telecommunication Systems. *IEEE Transactions on Vehicular Technology*, vol. 44 no. 3, August 1995, 543 – 553.
- [19] Mouly, M.; Pautet, M. B.: *The GSM System for Mobile Communication*. Published by the authors, 4, rue Elisée Reclus, F-91120 Palaiseau, France.
- [20] Reichl, P.; Kesdogan, D.; Junghärtchen, K.; Schuba, M.: Simulative Performance Evaluation of the Temporary Pseudonym Method for Protecting Location Information in GSM Networks. *Proceedings of the 10th International Conference for Computer Performance Evaluation TOOLS'98. Palma de Mallorca, Sept. 1998* (to appear in Springer LNCS).

