IDAMN: an Intrusion Detection Architecture for Mobile Networks[1]

Didier Samfat, Refik Molva

INSTITUT EURÉCOM
2229, route des Crêtes,
BP 193
06904 Sophia Antipolis,
FRANCE
Tel (+33) 93.00.26.26 Fax (+33) 93.00.26.27
{samfat, molva}@eurecom.fr

Abstract - **In this paper we present IDAMN, a distributed system whose main functionality is to track and detect mobile intruders in real-time. IDAMN includes two algorithms which model the behaviour of users in terms of both telephony activity and migration pattern. The main novelty of our architecture is its ability to perform intrusion detection in the visited location and within the duration of a typical call as opposed to existing designs that require the reporting of all call data to the home location in order to perform the actual detection. The algorithms and the components of IDAMN have been designed in order to minimize the overhead incurred in the fixed part of the cellular network.**

Keywords: Simulation, protocols, GSM, intrusion detection, network architecture, real-time, roaming, network facility fraud, subscription fraud.

1 Introduction

Internetworks of the future will allow and promote universal access. Users will be able to access the network at a multitude of access points separated by significant geographic distances and many administrative boundaries. This phenomenon has introduced new security issues compared to traditional fixed networks because of the lack of physical protection of the mobile network access points and of the transmission on the radio path. In order to protect a mobile network, two complementary security approaches can be considered: ***prevention*** and ***detection***.

The prevention approach consists in reducing the risk of threats by insuring that users respect the rules of usage of the network services. A well-known mechanism is ***authentication*** based on shared secrets [2]. The detection approach consists in looking for events indicating unusual activity on the network. The detection can be performed on-line and it is called ***intrusion detection***.

Several authentication protocols such as Global System for Mobile communications (GSM) [20], Digital European Cordless Telephone (DECT) [18] and Cellular Digital Packet Data

1. *The work described herein was part of a joint project funded by the IBM Zürich Research Laboratory*

(CDPD) [19] for personal communications systems have been proposed by standard Organizations. Other prevention based mechanisms, using different techniques, have also been proposed in the research literature [3],[4]. Independently of the approach adopted, the objectives are to counteract security threats such as masquerading and eavesdropping.

However, under certain situations, authentication mechanisms are not sufficient to protect the mobile network against threats such as the theft of the mobile unit, security holes in the software or hardware implementation, network facility and subscription fraud. Hence, if an adequate level of security is required, it is relevant to provide in addition to authentication a ***complementary*** security mechanism to mobile networks which is intrusion detection. Therefore, motivated by these problems, we made the specific design of a distributed Intrusion Detection Architecture for Mobile Networks (IDAMN).

This paper presents the functionalities of IDAMN as well as the intrusion detection algorithms used. We begin by briefly presenting the limitation of existing intrusion detection systems for both fixed and mobile networks. We then present in Section 3 the simulation platform which was used to test our prototype. In Section 4 we describe the different network entities of IDAMN allowing the monitoring of the telephony activity and the itineraries of the mobile users. The corresponding intrusion detection algorithms are described in Section 5 and Section 6 respectively. Finally, we demonstrate the efficiency of our algorithms and show that an implementation of IDAMN in a real network is feasible.

2 Background and Motivation

The first intrusion detection systems (IDS) were developed for the fixed networks. Their goal was to provide a sense of security while allowing computers and data networks to operate in an open mode. In fact, early IDSs were designed to monitor a single host by analysing audit trails provided by operating systems of computers [12], [13]. However, more recent IDS have been developed to accommodate the monitoring of hosts interconnected by a local area network. Examples of such systems are SRI's Intrusion Detection Expert System (IDES) [9], [11], Los Alamos National Laboratory's Network Anomaly Detection and Intrusion Reporter (NADIR) [8] and UC Davis' Distributed Intrusion Detection System (DIDS) [14]. A review of these IDS can be found in [5] and [6].

Even if such IDSs are efficient in their contexts, they can hardly be ported to a wide area network managing millions of subscribers. Moreover, they can not take into account the mobility of users who may be connected to large scale networks from different access points separated by significant geographic distances.

In the case of mobile networks, the first IDS was developed for the analog cellular network Advanced Mobile Phone System (AMPS) [16]. AMPS prevention mechanism is only based on ***identification*** and the open access to the radio gave intruders the opportunity to ***masquerade*** as a legitimate subscriber to make free calls. Moreover, as it was impossible for operators to add authentication mechanisms in the existing network, the only solution was therefore to build an IDS to avoid improper billing of legitimate subscribers. The basic idea of this system is to modify the existing database Home location Register HLR and the Visitor Location Register (VLR) in order to add some intrusion detection procedures. These procedures are in fact derived from algorithms based on the detection of credit card fraud.

The most recent system for mobile networks is the IDS developed by Digital Equipment Corporation [17]. The goal of this system is to provide intrusion detection mechanisms for analog

and for digital mobile networks. Hence, it is able to detect clones in analog mobile networks as well as to alleviate subscription fraud in GSM. This IDS relies on the ability of the Mobile Switching Centres (MSC) to temporarily record the call data generated by each mobile user. Then, the monitored information is transferred to a central site (Fraud Manager) for processing, and the verifications are performed «off-line».

In general, even if some existing commercial mobile networks contain authentication mechanisms, operators still want an external IDS to monitor their subscribers in order to alleviate unknown potential threats. Because the above mentioned systems need at least 3 days[1] to perform intrusion detection, it is relevant to build an efficient architecture while minimizing the overhead incurred in the mobile network. Moreover, the existing algorithms are based on thresholds which are delicate to choose as they usually depend on the human experience of the system.

In addition to avoid the aforementioned drawbacks of existing IDSs for mobile networks, we have taken into account the following design criteria:

- Minor modifications of the existing mobile network entities
- Fast algorithms to track intruders without having any *a priori* knowledge of the subscriber's behaviour
- Minimal transfer of information between IDAMN entities in order to detect an intrusion

In contrast to the existing systems that analyse the mobility of users by simply verifying the speed of mobile units, IDAMN benefits from a new algorithm for the verification of the user's itinerary pattern.

3 The simulation platform

The experimental platform is composed of IDAMN and of a Wireless NEtwork Simulator (WINES) which is a complex digital cellular network simulator implemented in accordance to the GSM technical specifications. However, IDAMN is not specific to GSM networks and can easily be adapted to other cellular networks; in that case the module converting signalling messages into statistical data must be adapted accordingly. Both WINES and IDAMN have been developed with a modular approach using OPNET software [26].

The reasons for choosing a GSM network simulator are twofold. Firstly, GSM does not provide intrusion detection services to operators despite the fact that *network facility* and *subscription fraud* have been encountered in this network. Secondly, the provision of a wide range of traffic generators spread over a wide geographic area is costly and difficult to perform. These problems can be overcome by the use of simulators which also present the advantage of exact repeatability of successive runs (useful during software implementation).

3.1 GSM specific procedures

WINES is able to simulate the whole network as well as the mobile stations at the protocol level. In other words, the resulting platform model simulates the Base Station Sub-system, the Network Sub-system and the Mobile Station (all GSM and IDAMN entities are independent software modules). All layer-3 messages exchanged by the different GSM entities, have the

1. *Depending of the actual location of the user*

same format as described in the GSM technical specifications. A full description of WINES can be found in [7]. The following procedures specific to digital cellular mobile networks have been implemented:

- ***Mobile Originating call*** (MO). The mobile user can make a call towards the Public Switching Telephony Network (PSTN); the frequency and duration of these calls are parametrizable as well.
- ***Mobile Terminating call*** (MT). Calls can also be initiated by the PSTN which asks the Gateway Mobile Switching Centre (MSC) for routing information.
- ***Handover***[1]. During a call the Base Station Controller (BSC) continuously analyses the measurements results sent by the Base Tranceiver Station (BTS) in order to decide on a potential handover. Depending on the situation, 4 types of handovers can be performed (*internal-handover*, *inter-BSC-handover*, *inter-MSC-handover*, *subsequent-handover*) as described in [22].
- ***Location updating*** (locup). This procedure is triggered by the MS each time its current location area changes [23]. If the old location area is under the control of the same MSC then an *inter-BSC-locup* procedure is performed. Otherwise, if the old and the new location area are managed by two distinct MSCs then an *inter-MSC-locup* procedure is performed.

During each simulation run, the GSM platform will generate signalling traffic data which will be collected by IDAMN in order to detect intrusive activities.

3.2 Traffic Generators

It is difficult to obtain audit data generated by real GSM users from operators who want to insure the confidentiality of their network parameters and the privacy of their subscribers. In order to generate realistic user data traffic within the network simulator, we made the specific design of a mobile stations ***Generator***.

The goal of a Generator is to simulate various mobile station behaviour models in order to test the different intrusion detection algorithms of IDAMN. In avoiding to prevaricate the detection results, the characteristics of the generated traffic must be as close as possible to the real traffic for a given GSM topology. Therefore, each parameter of a Generator is adjusted to activate a particular category of ***real*** GSM subscribers as presented in Table 1. This table describes GSM subscribers in terms of marketing[2] groups and was provided by a French operator. In other words, the values obtained from Table 1 were used to set the different parameters of the mobile station generators in the simulation platform.

Each Generator is able to simulate a group of 1000 mobile stations with the same behaviour. The frequency and the duration of MO calls are parametrizable and the peak hours (busy period during the day) can also be defined. For each user, a behaviour scenario is established at the beginning of the simulation depending on two statistical distributions: the duration of calls follows an exponential distribution whereas the arrival of calls follows a Poisson distribution. These assumptions have been shown to be reasonably accurate for mobile telephone networks. In addition to the telephony activity, the parameters corresponding to the mobility of users can be defined (number of location updates and location areas traversed).

---

1. *Changing of cell without interrupting the call.*
2. *This classification of GSM subscribers was obtained from a french operator*

Type of User	DOMESTIC	CORPORATE	BUSINESS	ROAMER
Daily Usage [a]	**10-15**	30-40	15-20	30 +
Calls per Week	12-16	20-25	12-15	25 +
National Call	95%	75%	99%	95%
International Call	5%	25%	1%	5%
Average Duration[b]	< 5	10	< 10	10 +
Call Time	Off peak	Peak Time	business Hours	Any
Destination Call	Mostly local	Nation Wide	Local	International
Origin of Call	Home Cell	Any MSC/BSC	Same MSC	All
Type of Call	MTL[c]	MTL and **MTM**[d]	MTL	MTL and MTM

Table 1 : Classification of GSM subscribers

a. Value in minutes
b. Value in minutes
c. Mobile to Land
d. Mobile to Mobile

A simulation run can represent the behaviour of the network during one day. Therefore, a single Generator is able to simulate a population of 400 mobile stations during working days or week-ends.

4 Functionalities of IDAMN

The challenge was to model IDAMN in order to detect an intruder «***on-line***» while minimizing the overhead incurred in the cellular network. For this purpose, we have defined a set of analysis functions which are spread over the whole network. In fact, these functions are performed by different distributed entities of IDAMN in order to monitor mobile users.

The main novelty of our architecture is its ability to perform intrusion detection in the visited location and within the duration of a typical call as opposed to existing designs that require the reporting of all call data to the home location in order to perform the actual detection. In other words, the performance of IDAMN rely on the transfer of statistical information of small size instead of centralizing large amounts of audit data like in existing IDSs for mobile networks [16], [17].

4.1 Multi-Level Intrusion Detection

We have defined three different levels of analysis which can be summarized as follows:

- ***Level 1 - velocity and clone verification:*** this consists of a fast intrusion detection by verifying the speed of a mobile user or existing clones (the same user being active in two different parts of the network at the same time).
- ***Level 2 - component wise verification:*** the system measures the impact of the subscriber behaviour on the different GSM entities. For instance high activity of a switch in an area with a low density of inhabitants, may be a symptom of intrusion.

- ***Level 3 - intrusion detection per user***: this is the most significant intrusion detection analysis as the resulting procedure evaluates every deviation from the normal user's normal behaviour profile or «signature».

In the case of the level-3 analysis the basic idea of detecting an intruder relies on the system's ability to learn the normal behaviour of the subscriber by creating a ***user profile***. The profiles of IDAMN are designed to require a minimum amount of storage for historical data per mobile user and yet require relevant information that can efficiently be used during the detection process. An intruder impersonating the real subscriber will have a different behaviour and thus will generate a significant deviation from the standard profile [9].

In the case of GSM, the signature of the user is defined by 3 profiles: a mobility-profile, an activity-profile and a speech-profile[1]. Each profile will help in raising different intrusion alarms that a rule based system will analyse in order to give the final decision. Therefore, forbidding a mobile user to lend his mobile station becomes a mandatory security policy in this context.

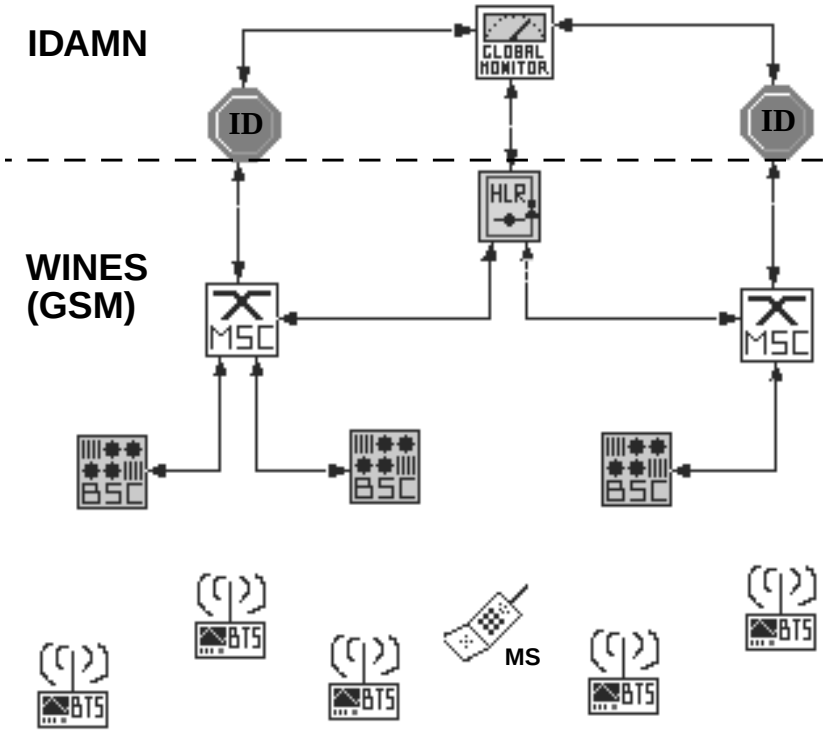

Figure 1 : Portion of the Global Platform

4.2 IDAMN Network Entities

The relevant information for modelling the user's behaviour as well as detecting an intrusion are distributed over the various entities of the GSM Network Sub-System. In order to minimize the computational overhead incurred in both IDAMN and GSM, the procedures addressing the different level of intrusion detection are spread over two dedicated machines: the ***Global Monitor*** (GM) and the ***Intrusion Detector*** (ID) as depicted in Figure 1.

1. *A speech verification algorithm is envisioned but has not been developed yet.*

The main purpose of the GM is to manage and to save over a long period the profiles of the subscribers. Moreover, the GM is in charge of executing the procedures corresponding to level 1 and level 2 of intrusion detection. The results of these 2 first analyses are processed by a rule based system which generates a *global report* to be passed to the ID.

The ID performs level 1 (detection of a clone at the location area level) and level 3 of intrusion detection. The ID is directly connected to the anchor[1] MSC and makes a copy of the relevant signalling messages generated by the mobile users. Depending on the nature of the messages, the ID initializes a set of statistical variables and requests the activity and mobility profiles from the GM as well as the global report concerning the mobile user. Then, the ID is able to detect a potential intrusion by comparing the variables to the received profiles according to some thresholds.

The GM is able to access the HLR via the GSM standard protocol MAP/C in contrast to the ID which can retrieve information from the VLR using the MAP/B protocol [24] if required. Moreover, a single ID is associated with each MSC allowing IDAMN to have a complete view on the mobile user in order to analyse his ***telephony activity*** and his ***migration pattern*** as discussed in the following sections.

5 Monitoring the Telephony Activity

The telephony activity of a mobile user is defined by two statistical vectors: the ***Call Vector*** and the ***Session Vector***. The values of these vectors are initialized by the ID upon receiving signalling messages from the anchor MSC as depicted in Figure 2.

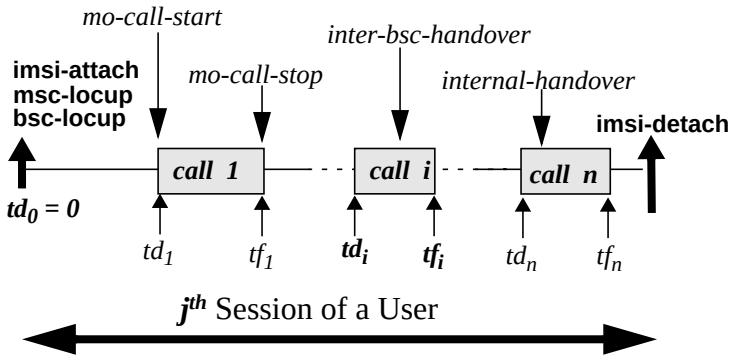

Figure 2 : Call Vectors initialized with GSM signalling messages

The Call Vector $\vec{V}_i^c$ is a local magnitude which models the i^{th} MO. We denote by td_i and f_i respectively the time when the MS triggered a *mo-call-start* and a *mo-call-stop* procedure. Upon receiving the GSM call management messages, the ID computes a statistical call vector over the duration of an outgoing call ($f_i - td_i$) , number of handovers performed[2] (nh_i) and inactivity between two calls ($d_i - tf_{i-1}$). The expression of $\vec{V}_i^c$ is as follows:

1. *In GSM terminology, this is an MSC managing the MS at the beginning of a call*
2. *Each time a handover procedure is performed, nh_i is increased by one*

$$\vec{}^{c}_{i} = \begin{bmatrix} td_i - tf_{i-1} \\ tf_i - td_i \\ nh_i \end{bmatrix} \quad (1)$$

The Session Vector $\vec{V}^s_j$ is a global magnitude which models the j^{th} session of the user in terms of network connection duration (D_j) total number of calls (Nc_j total duration of calls (Dc_j) and the total number of handovers Nh_j. The expression of $\vec{V}^s_j$ is the following:

$$\vec{}^{s}_{j} = \begin{bmatrix} D_j \\ Dc_j \\ Nh_j \\ Nc_j \end{bmatrix} \quad (2)$$

The high-speed intrusion detection algorithm which is able to interpret these vectors is based on the IDES approach [10]. However, our algorithm takes into account the specific context of wide cellular networks managing millions of subscribers such as GSM, whilst minimizing the computation overhead and the amount of data transferred between IDAMN entities.

In doing so, the telephony activity profile of a user is updated in a recursive way each time a new vector[1] $\vec{V}_n$ (i.e. $\vec{V}^s_n$ or $\vec{V}^c_n$) is initialized by the ID. Moreover, before incorporating a new vector $\vec{V}_{n+1}$ into the profile P_A, which is composed of a mean vector $\vec{M}_n$ and a covariance matrix C_n, we first compute $\vec{M}_n$ and C_n by multiplying each observed vector $\vec{V}_i$ by the following weight:

$$w_i = \frac{\alpha^{n-i}}{\sum_{i=1}^{n} \alpha^{n-i}} \quad (3)$$

where α is the forgetting factor ($0 < \alpha < 1$). This method creates a sliding time window for the observed vector which has the effect to favour the recent behaviour of the user. Therefore, if P_A has been previously computed over first *n* vectors (with *n* smaller than 30), the ID updates the mean vector as follows:

$$\vec{M}_{n+1} = \frac{1}{1-\alpha^{n+1}}(\alpha(1-\alpha^n)\vec{M}_n + (1-\alpha)\vec{V}_{n+1}) \quad (4)$$

and the covariance matrix is computed as follows:

$$\begin{aligned} C_{n+1} = {} & \frac{\alpha(1-\alpha^{n-1})}{1-\alpha^n} C_n + \frac{1-\alpha}{2}(\vec{M}_n - \vec{M}_{n+1})(\vec{M}_n - \vec{M}_{n+1})^t \\ & + \frac{1-\alpha^2}{2\alpha(1-\alpha^n)}(\vec{V}_{n+1} - \vec{M}_{n+1})(\vec{V}_{n+1} - \vec{M}_{n+1})^t \end{aligned} \quad (5)$$

1. *Whether it is a call vector or a session vector*

where *t* denotes the matrix transposition operation. Hence, the GM will not have to store all observed vectors in its database as the profile to be saved is $P_A = \left\{\overrightarrow{M_n}; C_n; \alpha\right\}$.

Once P_A has been generated, detecting an abnormal activity consists in computing the distance between a new vector $\vec{V}_{n+1}$ and the mean vector which is a function of the cluster of the *n* previous vectors. If this distance is smaller than a threshold S_{max} then, the new $\vec{V}_{n+1}$ is considered to be normal. Otherwise, the following inequation denotes the detection test for the abnormal case:

$$(\vec{V}_{n+1} - \vec{M}_n)^t C_n^{-1}((\vec{V}_{n+1} - \vec{M}_n) \geq S_{max}^2) \quad (6)$$

where C_n^{-1} represents the inverse matrix of C_n. We performed several experiments using different kinds of thresholds aiming to maximize the detection rate and to minimize the false alarm rate. Consequently, we obtain the best results with the following decision criterion:

$$S_{max} = max\left\{(\vec{V}_i - \overrightarrow{M_n})^t C_n^{-1}(\vec{V}_i - \overrightarrow{M_n})\right\}(for \quad i \in [1 \ldots n]) \quad (7)$$

The threshold S_{max} is not *a priori* fixed but is computed over the *n* first observed vectors for each user. Other thresholds have been used but the false alarm rate was high (between 20 to 50%) even if all intrusion scenarios were detected.

Category of user	Call false alarm rate	Call detection rate	Session false alarm rate	Session detection rate
DOMESTIC	1%	67 to 100%	2%	80 to 100%
BUSINESS	1%	88 to 100%	2%	90 to 100%
CORPORATE	1%	60 to 100%	5%	87 to 100%
ROAMER	2%	82 to 100%	1%	95 to 100%

Table 2 : False Alarm and Detection rates

For each mobile user belonging to a specific category of subscriber, 300 intrusive session vectors were tested (one session representing one day of connection to the network). Almost[1] every intrusive session was detected (at least 80%) with a false alarm rate varying from 2% to 5% as shown in Table 5. Much better performance was obtained with the call vector analysis (600 to 2000 vectors tested per user) as the false alarm rates were under 2% with a minimum detection rate of 60%. In every case, if a high intrusive activity is simulated, 100% of the matching vectors are detected; this corresponds to an intruder trying to misuse the network services extensively as already noticed in existing analog mobile networks.

6 Monitoring Users' Roaming

In addition to the telephony activity analysis, IDAMN includes a detection mechanism based on the monitoring of the users' travelling in the network. This mechanism It takes into account

1. *Depending on the category of user (domestic, corporate etc.)*

the areas visited by a subscriber as well as the most frequent itineraries followed. A mobility behaviour profile based on a graph model (with transition probabilities) is established depending on the frequency of location area crossings of the mobile user. This mobility profile is updated when the location updating procedures are performed (i.e, *inter-MSC-locup inter-BSC-locup*).

We assume that the migration of a mobile user is not completely random. For instance, most of the time a BUSINESS user will roam from his house to his office through a reasonable number of location areas and well defined cells. For the sake of simplicity, in the rest of the paper we denote by «cell» a location area.

Then, we associate to each cell a state, and compute the state transition probability. Depending on the itinerary[1] observed from the location update signalling messages, the probabilities of the different itineraries are calculated in order to give more importance to those the most followed. In other words, the transition probability going from a state *i* to a state *j* is a function of the number of times the mobile user has effectively traversed cell *i* to cell *j*.

Let X_n denote the stochastic variable indicating the cell on which the mobile station is located after the n^{th} location update. Consequently, if $X_n = 0$ the MS is considered to be disconnected and if $X_n = i$ the MS is camping on cell *i*. Therefore, given that at the previous ((n-1)'th) location update the mobile was camping on cell i, the transition probability from cell *i* to *j* is:

$$\begin{aligned} {}_{ij} &= p\langle X_n = j \mid X_{n-1} = i\rangle = \frac{n_{ij}}{n_{i_\circ}} \\ n_{i_\circ} &= \sum_k n_{ik} \end{aligned} \tag{8}$$

where n_{ij} is the number of times a user crossed cell *i* to cell *j* and $n_{i_\circ}$ is the total number of output occurrences of cell *i*. At the initial step, we consider that the mobile user can go in all directions with the same probability[2]. However, the training period will give higher probabilities to the most frequent itineraries depending on the migration habits of the user. Hence, the mobility profile of a user is a graph P_M where each cell is a state and each transition a probability as defined by Equation (8). In the example of Figure 3, the user crossed cells 2, 3 and 4 *n* times whereas cells 1, 3 and 5 were traversed only once.

The detection process consists of comparing the current itinerary in progress (called candidate itinerary) to the profile P_M in order to measure a strong deviation from that profile. Let $\hat{C}$ be the candidate itinerary composed of *k* cells; for each cell *i* of $\hat{C}$ we compute the average probability $\bar{p}_i$ as follows:

$$\bar{p}_i = \frac{1}{O_i} \cdot \sum_{\substack{j=1 \\ j \neq i}}^{n} p_{ij} = \frac{1 - p_i}{O_i} \qquad i \in [0 \ldots k] \tag{9}$$

where O_i is the total number of outgoing transitions of cell *i*.

1. *An itinerary is composed of successive cells or location area*
2. *All probabilities are initialized to zero*

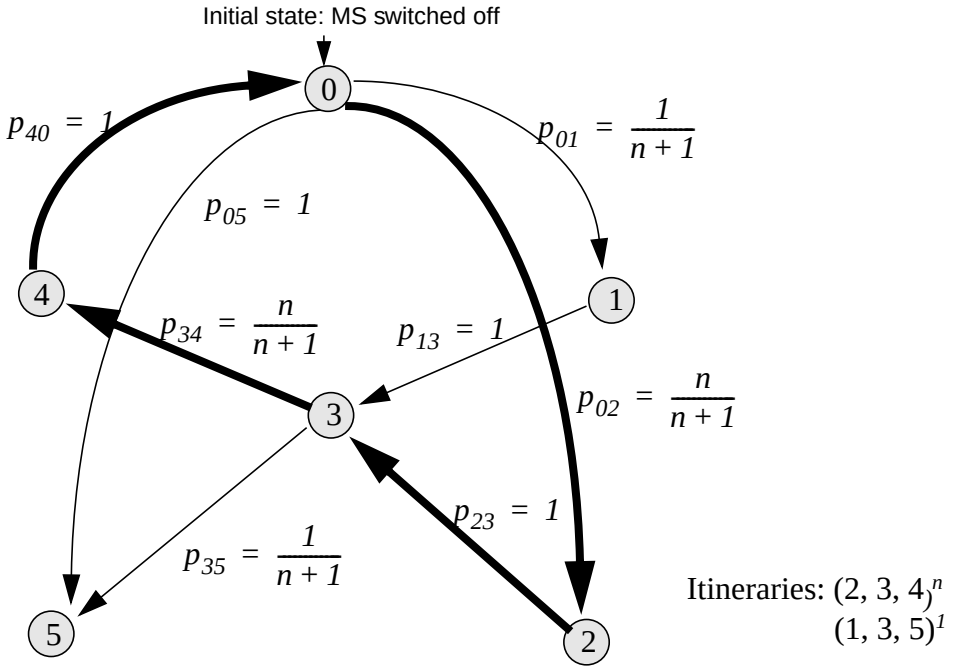

Figure 3 : Example of a Mobility Profile

The ultimate detection test is performed by analysing a doubt factor V_d ($V_d \geq 0$). Moving from cell i to cell j will be considered as abnormal if $p_{ij} \geq \bar{p}_i$. In that case, the doubt factor is reduced as follows: $_d = KV_d$ (with $0 < k < 1$). Otherwise, if $p_{ij} < \bar{p}_i$ then V_d is augmented with a positive value depending on the following situations:.

- If the cell has already been crossed then the doubt factor is increased by H as follows: $_d = V_d + H$.
- If the mobile unit arrives in a new cell for the first time V_d is incremented as follows: $V_d = V_d + \frac{F}{2^l}$,

where l represents the number of cells composing $\hat{C}$ which have never been encountered during earlier update of P_M.

As $\lim_{n \to \infty} \sum_{i=1}^{n} \frac{F}{2^i} = 2F$, we consider that the itinerary is abnormal if $V_d > 2F$. This decision criterion is tolerant to potential new itineraries of the legitimate user who may roam into visited domain instead of his usual home domain. This algorithm performs intrusion detection on-line as each time the mobile user performs a location update procedure it is possible to know whether or not the current cell traversed is normal.

In fact, parameters F and H depend on the average length of the user's itineraries. For each mobile user, 300 abnormal itineraries were tested over a population of 400 simulated users. The false alarm rate is lower than 15% with a detection rate varying from 65% to 100% as shown in Table 5.

Category of user	False alarm rate	Detection rate
DOMESTIC	1 to 2%	90 to 100%
BUSINESS	3 to 4%	75 to 90%
CORPORATE	4 to 6%	65 to 90%
ROAMER	5 to 7%	65 to 80%

Table 3 : False Alarm and Detection rates

7 IDAMN in Practice

Intrusion alarms are raised when the detector notices a strong deviation of the mobile user behaviour compared to the mobility and activity profiles. Next, these alarms are analysed by a rule based system which gives the final decision. If an intrusion is detected IDAMN can signal the network to activate a call bearing procedure or to disconnect the user.

The idea of comparing profiles of users with current session data is not exempt from potential problems. An intruder impersonating the subscriber can gradually modify his behaviour in order to escape detection. However, this implies that the attacker knows *a priori* the telephony activity as well as the migration pattern of the user. In a real large scale wireless network, it will be difficult for an attacker to obtain ***all*** audit data generated by the mobile user if a minimal service of privacy is provided by the means of encryption[1]. On the other hand, if the mobile network does not benefit from authentication and confidentiality mechanisms, then IDAMN must be provided with additional algorithms. For instance, intrusion detection algorithms based on speech verification or dialled number analysis will be an issue. Nevertheless, if we assume that an intruder having succeeded to obtain fraudulent access, his goal will be to use extensively the network services and thus will be automatically be detected considering the current version of IDAMN.

In order to demonstrate that implementing IDAMN in a real mobile network is feasible, we need to prove that the architecture is scalable and the algorithms have a reasonable complexity. Therefore, we denote by ***flop*** (floating point operation) an operation which require one multiplication and one addition: $1\ flop = S_{k-1} + a_k b_k$.

Thus, the number of flops needed to update the activity profile $P_A = \left\{ \overrightarrow{M_n}; C_n; \alpha \right\}$ is $(m^2 + 3m)/2$ where m is the dimension of the statistical vector. In comparison, without the recursive formula, the total number of flops needed to update P_A will be around $3m^2$. In the case of the intrusion detection test, we can also compute C_n^{-1} in a recursive way and hence, the number of flops needed to perform detection is $2m$.

With a DSP of the TMS 320 family [25], a floating point number is coded in 4 bytes and an integer in 2 bytes. Therefore, the global monitor of IDAMN will need $\left(\frac{m^2 + 3m}{2} + 1\right) \times 4$ bytes

1. *After being authenticated by the network the user shares a session key for the encryption of the signalling and user data.*

for a user's profile (P_A). The DSP TMS 320 is able to execute 33 million flops per second. In our case, we suppose that we need only 1 μs to execute 1 flop. Table 4 summarizes the characteristics of our algorithm.

Characteristics	Call Vector (m=3)	Session Vector (m=4)
Time to update a profile	9 µs	14 µs
Time for detection	6 µs	8 µs
Size of a profile	40 bytes	60 bytes

Table 4 : Characteristics of the activity analysis

If the profile P_M of a user is a graph composed of *m* states (or cells) then the complexity for creating or updating of P_M is m^2. The computation of $\bar{p}_i$ will need around *m* flops and the total number of operations to perform detection is $m^2 + m$.

As opposed to the telephony activity algorithm, the memory size to store P_M varies from a category of user to the other. For instance, a DOMESTIC user is rather static and the size of his mobility profile is smaller than the ROAMER user profile. Table 5 presents the characteristics of our algorithm.

Category of user	Number of cells[a]	Matrix size[b]	Detection Duration
DOMESTIC	4	32	16 µs
BUSINESS	10	200	100 µs
CORPORATE	16	512	256 µs
ROAMER	24	1152	576 µs

Table 5 : Characteristics of the migration analysis

a.In the graph
b.in bytes

The values shown in this table have been computed by considering the maximum number of cells traversed by each category of user. This algorithm might require more resources than the vector test algorithms but the time for the detection is also in the order of μs. Hence, the total size for both the activity and migration profile for one user is at most 1,5 Kb.

A «***real world***» implementation of IDAMN can be envisioned. The planning of IDAMN depends on the topology of the mobile network and in our case on GSM. In general, a HLR manages 500 000 subscribers and a MSC 40 000 subscribers [1]. Therefore, if the network comprises 1 000 000 mobile users[1], IDAMN will be composed of 2 GMs and 25 IDs which are interconnected with a 64 Kbits/s leased line and each GM managing at most 13 IDs.

1. *For instance, 800 000 subscribers are currently using the GSM network of the french operator France Telecom*

According to [1], during peak hours a MSC receives in average 24 MO calls per second but for demonstration, we consider the worst case of 40 MO calls per second. If we assume that this load is the same for the 13 MSCs, then each ID sends in 65 ms $40 \times 100 = 4$ Kbits of profile requests[1] to one GM. Upon receiving these profile requests, the GM must search for the 40 profiles (1 ms needed per user profile[2]) in order to transmit them to the corresponding ID. The time needed to perform these operations is around $1,5 \times 40 / 64 \approx 0,9\ s$. Hence, because an ID takes at most 600 μs to detect an abnormal behaviour, all intrusion detection algorithms can be performed in less than 1 second.

Note that the user profiles are transmitted only once between the GM and an ID. Upon receiving new activity or mobility data, the same ID will be able to perform detection immediately without contacting the GM. Moreover, IDAMN does not save any call detail records because if the call is normal then the ID updates the profile and the call data is dropped. This has the advantage to minimize the amount of storage compared to centralized IDSs which have to save several days of call detail records (cellular billing information) in order to allow the fraud investigator to view the call activity of the mobile user.

8 Conclusion

This paper presented IDAMN a distributed intrusion detection system for cellular networks. IDAMN is able to have a complete overview on the mobile user telephony activity and mobility behaviour. The main novelty of our architecture is its ability to transfer a small user profile near the location of the serving MSC as opposed to existing systems which require the reporting of large amount of audit data to a central monitoring entity. The algorithms have been designed in order to create and update the user profiles in real time without having *a priori* knowledge of the behaviour of the user. Moreover, an alarm can be raised within the duration of a typical call. Preliminary testing of the components of IDAMN are promising as we can raise an alarm in less than one second with an intrusion detection rate varying from 70% to 100% whereas the false alarm rate is lower than 5%. Future work includes and enhancement of IDAMN's functionalities by developing speech and dialled number verification algorithms.

9 Acknowledgments

We are indebted to Jacques Labetoulle and Dirk Slock for their insightful comments on previous versions of the intrusion detection algorithms. We would like also to thank Raymond Knopp for his assistance during the preparation of this article and Christian Bonnet for the implementation of the GSM simulator WINES.

References

[1] M. Mouly, M.B. Pautet, «The GSM System for Mobile Communications», ISBN 2-9507190-0-7, 1993.

[2] R. Molva, D. Samfat and G. Tsudik, «Authentication of Mobile Users», IEEE Network Magazine, Special Issue on Mobile Communications, March/April 1994.

[3] Ashar Aziz and Whitfield Diffie, «Privacy and Authentication for Wireless Local Area Networks», IEEE Personal Communication, First Quarter 94.

[4] M. Beller, L. Cheng and Y. Yacobi, «Privacy and Authentication on a Portable Communication System»,

1. *In the protocol we have implemented, the size of the profile request is 100 bits.*
2. *This is a time needed to a VLR to retrieve information in the database.*

IEEE Journal on Selected Area in Communication, Aug. 1993.

[5] B. Mukherjee, L.T. Herbelein, K.N. Levitt, Network Intrusion Detection, IEEE Network Magazine, May/ June 1994, VOL. 8 No 3.

[6] G.B. White, E.A. Fish and U.W. White, «Cooperating Security Managers: A Peer-Based intrusion Detection System», IEEE Network Magazine, January/February 1996.

[7] D. Samfat, V. Devernay, C. Bonnet, «A GSM Simulation Platform for Intrusion Detection», Proceedings of ICC'95, Seattle, June 1995.

[8] J. Hochberg et al., « NADIR: an Automated System for Detecting Network Intrusion and Missue», Computer and Security, vol. 12, no 3, May 1993.

[9] T.F. Lunt, "IDES: An Intelligent System for Detecting Intruders", Proceedings of the Symposium: Computer Security, Threat and Countermeasures, Rome, Italy, November 1990.

[10] Theresa F. Lunt, «Using Statistics to Track Intruders», Proceedings of the Joint Statistical Meetings of the American Statistical Association, August 1990.

[11] T.F. Lunt, «A Real-time Intrusion Detection Expert System (IDES)», Interim Progress Report, Project 6784, SRI International, May 1990.

[12] K. Ilgun, P.A Poras, R.A Kemmerer, «USTAT a Real-Time Intrusion Detection System For Unix», SRI Intrusion Detection Workshop, Paolo Alto, CA Feb. 1989

[13] S.E. Smaha, «Haystack: An Intrusion Detection System», Proceedings of the IEEE Fourth Aerospace Computer Security Applications Conference, Orlando, FL, Dec. 1988

[14] Steven R. Snapp et al. , "DIDS (Distributed Intrusion Detection System) - Motivation, Architecture and an Early Prototype", Proceedings of the 14th National Computer Security Conference, Washington, DC, Oct 91.

[15] L.T. Hebberlein, K.N Levitt and B. Mukherjee, "An Intrusion-Detection System for Large-Scale Networks", Proceedings of the 15th National Computer Security Conference, Baltimore, MD, October 1992.

[16] Rob Mechaley and Kirk Calson, «Parameters for Fraud Management Using Network Based Techniques», Technical Report 45.2.II.2, Mc Caw Cellular Communications, Inc, 1991.

[17] Evan J. Davies, Henrik Nordin, Sharon Hershon, Kathy Gallup, «Design Overview for Fraud Detection and Analysis System», Technical Report, Digital Equipment Corporation, August 1992.

[18] ETS 300 175-7, «Part 7: Security Features», DECT common Interface, ETSI, Oct. 1992.

[19] «Cellular Digital Packet Data (CDPD)», System Specifications, Release 1.0, July 1993

[20] GSM Recommandation 3.20, «Security Related Network Function», ETSI Standard, Feb. 1992.

[21] GSM Recommendation 3.05, Signalling Requirements Relating to Routing of calls to Mobile Subscribers, ETSI Standard, Feb. 1992.

[22] GSM Recommendation 3.09, Handover Procedures, ETSI Standard, Feb. 1992.

[23] GSM Recommendation 3.12, Location Registration Procedures, ETSI Standard, Feb. 1992.

[24] GSM Recommendation 9.02, MAP Specification, ETSI Standard, Feb. 199.

[25] Panos Papamichalis, "Digital Signal Processing Application with the TMS320 Family", Application Book, Texas Instruments 1990.

[26] OPNET Network Simulation Software, MIL 3, Inc. 3400 International Drive, Washigton, DC 20008, Release 2.4.A.

IDAMN: an Intrusion Detection Architecture for Mobile Networks[1]

Didier Samfat, Refik Molva

INSTITUT EURÉCOM
2229, route des Crêtes,
BP 193
06904 Sophia Antipolis,
FRANCE
Tel (+33) 93.00.26.26 Fax (+33) 93.00.26.27
{samfat, molva}@eurecom.fr

Abstract - **In this paper we present IDAMN, a distributed system whose main functionality is to track and detect mobile intruders in real-time. IDAMN includes two algorithms which model the behaviour of users in terms of both telephony activity and migration pattern. The main novelty of our architecture is its ability to perform intrusion detection in the visited location and within the duration of a typical call as opposed to existing designs that require the reporting of all call data to the home location in order to perform the actual detection. The algorithms and the components of IDAMN have been designed in order to minimize the overhead incurred in the fixed part of the cellular network.**

Keywords: Simulation, protocols, GSM, intrusion detection, network architecture, real-time, roaming, network facility fraud, subscription fraud.

1 Introduction

Internetworks of the future will allow and promote universal access. Users will be able to access the network at a multitude of access points separated by significant geographic distances and many administrative boundaries. This phenomenon has introduced new security issues compared to traditional fixed networks because of the lack of physical protection of the mobile network access points and of the transmission on the radio path. In order to protect a mobile network, two complementary security approaches can be considered: ***prevention*** and ***detection***.

The prevention approach consists in reducing the risk of threats by insuring that users respect the rules of usage of the network services. A well-known mechanism is ***authentication*** based on shared secrets [2]. The detection approach consists in looking for events indicating unusual activity on the network. The detection can be performed on-line and it is called ***intrusion detection***.

Several authentication protocols such as Global System for Mobile communications (GSM) [20], Digital European Cordless Telephone (DECT) [18] and Cellular Digital Packet Data

1. *The work described herein was part of a joint project funded by the IBM Zürich Research Laboratory*

(CDPD) [19] for personal communications systems have been proposed by standard Organizations. Other prevention based mechanisms, using different techniques, have also been proposed in the research literature [3],[4]. Independently of the approach adopted, the objectives are to counteract security threats such as masquerading and eavesdropping.

However, under certain situations, authentication mechanisms are not sufficient to protect the mobile network against threats such as the theft of the mobile unit, security holes in the software or hardware implementation, network facility and subscription fraud. Hence, if an adequate level of security is required, it is relevant to provide in addition to authentication a ***complementary*** security mechanism to mobile networks which is intrusion detection. Therefore, motivated by these problems, we made the specific design of a distributed Intrusion Detection Architecture for Mobile Networks (IDAMN).

This paper presents the functionalities of IDAMN as well as the intrusion detection algorithms used. We begin by briefly presenting the limitation of existing intrusion detection systems for both fixed and mobile networks. We then present in Section 3 the simulation platform which was used to test our prototype. In Section 4 we describe the different network entities of IDAMN allowing the monitoring of the telephony activity and the itineraries of the mobile users. The corresponding intrusion detection algorithms are described in Section 5 and Section 6 respectively. Finally, we demonstrate the efficiency of our algorithms and show that an implementation of IDAMN in a real network is feasible.

2 Background and Motivation

The first intrusion detection systems (IDS) were developed for the fixed networks. Their goal was to provide a sense of security while allowing computers and data networks to operate in an open mode. In fact, early IDSs were designed to monitor a single host by analysing audit trails provided by operating systems of computers [12], [13]. However, more recent IDS have been developed to accommodate the monitoring of hosts interconnected by a local area network. Examples of such systems are SRI's Intrusion Detection Expert System (IDES) [9], [11], Los Alamos National Laboratory's Network Anomaly Detection and Intrusion Reporter (NADIR) [8] and UC Davis' Distributed Intrusion Detection System (DIDS) [14]. A review of these IDS can be found in [5] and [6].

Even if such IDSs are efficient in their contexts, they can hardly be ported to a wide area network managing millions of subscribers. Moreover, they can not take into account the mobility of users who may be connected to large scale networks from different access points separated by significant geographic distances.

In the case of mobile networks, the first IDS was developed for the analog cellular network Advanced Mobile Phone System (AMPS) [16]. AMPS prevention mechanism is only based on ***identification*** and the open access to the radio gave intruders the opportunity to ***masquerade*** as a legitimate subscriber to make free calls. Moreover, as it was impossible for operators to add authentication mechanisms in the existing network, the only solution was therefore to build an IDS to avoid improper billing of legitimate subscribers. The basic idea of this system is to modify the existing database Home location Register HLR and the Visitor Location Register (VLR) in order to add some intrusion detection procedures. These procedures are in fact derived from algorithms based on the detection of credit card fraud.

The most recent system for mobile networks is the IDS developed by Digital Equipment Corporation [17]. The goal of this system is to provide intrusion detection mechanisms for analog

and for digital mobile networks. Hence, it is able to detect clones in analog mobile networks as well as to alleviate subscription fraud in GSM. This IDS relies on the ability of the Mobile Switching Centres (MSC) to temporarily record the call data generated by each mobile user. Then, the monitored information is transferred to a central site (Fraud Manager) for processing, and the verifications are performed «off-line».

In general, even if some existing commercial mobile networks contain authentication mechanisms, operators still want an external IDS to monitor their subscribers in order to alleviate unknown potential threats. Because the above mentioned systems need at least 3 days[1] to perform intrusion detection, it is relevant to build an efficient architecture while minimizing the overhead incurred in the mobile network. Moreover, the existing algorithms are based on thresholds which are delicate to choose as they usually depend on the human experience of the system.

In addition to avoid the aforementioned drawbacks of existing IDSs for mobile networks, we have taken into account the following design criteria:

- Minor modifications of the existing mobile network entities
- Fast algorithms to track intruders without having any *a priori* knowledge of the subscriber's behaviour
- Minimal transfer of information between IDAMN entities in order to detect an intrusion

In contrast to the existing systems that analyse the mobility of users by simply verifying the speed of mobile units, IDAMN benefits from a new algorithm for the verification of the user's itinerary pattern.

3 The simulation platform

The experimental platform is composed of IDAMN and of a Wireless NEtwork Simulator (WINES) which is a complex digital cellular network simulator implemented in accordance to the GSM technical specifications. However, IDAMN is not specific to GSM networks and can easily be adapted to other cellular networks; in that case the module converting signalling messages into statistical data must be adapted accordingly. Both WINES and IDAMN have been developed with a modular approach using OPNET software [26].

The reasons for choosing a GSM network simulator are twofold. Firstly, GSM does not provide intrusion detection services to operators despite the fact that *network facility* and *subscription fraud* have been encountered in this network. Secondly, the provision of a wide range of traffic generators spread over a wide geographic area is costly and difficult to perform. These problems can be overcome by the use of simulators which also present the advantage of exact repeatability of successive runs (useful during software implementation).

3.1 GSM specific procedures

WINES is able to simulate the whole network as well as the mobile stations at the protocol level. In other words, the resulting platform model simulates the Base Station Sub-system, the Network Sub-system and the Mobile Station (all GSM and IDAMN entities are independent software modules). All layer-3 messages exchanged by the different GSM entities, have the

1. *Depending of the actual location of the user*

same format as described in the GSM technical specifications. A full description of WINES can be found in [7]. The following procedures specific to digital cellular mobile networks have been implemented:

- ***Mobile Originating call*** (MO). The mobile user can make a call towards the Public Switching Telephony Network (PSTN); the frequency and duration of these calls are parametrizable as well.
- ***Mobile Terminating call*** (MT). Calls can also be initiated by the PSTN which asks the Gateway Mobile Switching Centre (MSC) for routing information.
- ***Handover***[1]. During a call the Base Station Controller (BSC) continuously analyses the measurements results sent by the Base Tranceiver Station (BTS) in order to decide on a potential handover. Depending on the situation, 4 types of handovers can be performed (*internal-handover*, *inter-BSC-handover*, *inter-MSC-handover*, *subsequent-handover*) as described in [22].
- ***Location updating*** (locup). This procedure is triggered by the MS each time its current location area changes [23]. If the old location area is under the control of the same MSC then an *inter-BSC-locup* procedure is performed. Otherwise, if the old and the new location area are managed by two distinct MSCs then an *inter-MSC-locup* procedure is performed.

During each simulation run, the GSM platform will generate signalling traffic data which will be collected by IDAMN in order to detect intrusive activities.

3.2 Traffic Generators

It is difficult to obtain audit data generated by real GSM users from operators who want to insure the confidentiality of their network parameters and the privacy of their subscribers. In order to generate realistic user data traffic within the network simulator, we made the specific design of a mobile stations ***Generator***.

The goal of a Generator is to simulate various mobile station behaviour models in order to test the different intrusion detection algorithms of IDAMN. In avoiding to prevaricate the detection results, the characteristics of the generated traffic must be as close as possible to the real traffic for a given GSM topology. Therefore, each parameter of a Generator is adjusted to activate a particular category of ***real*** GSM subscribers as presented in Table 1. This table describes GSM subscribers in terms of marketing[2] groups and was provided by a French operator. In other words, the values obtained from Table 1 were used to set the different parameters of the mobile station generators in the simulation platform.

Each Generator is able to simulate a group of 1000 mobile stations with the same behaviour. The frequency and the duration of MO calls are parametrizable and the peak hours (busy period during the day) can also be defined. For each user, a behaviour scenario is established at the beginning of the simulation depending on two statistical distributions: the duration of calls follows an exponential distribution whereas the arrival of calls follows a Poisson distribution. These assumptions have been shown to be reasonably accurate for mobile telephone networks. In addition to the telephony activity, the parameters corresponding to the mobility of users can be defined (number of location updates and location areas traversed).

---

1. *Changing of cell without interrupting the call.*
2. *This classification of GSM subscribers was obtained from a french operator*

Type of User	DOMESTIC	CORPORATE	BUSINESS	ROAMER
Daily Usage [a]	**10-15**	30-40	15-20	30 +
Calls per Week	12-16	20-25	12-15	25 +
National Call	95%	75%	99%	95%
International Call	5%	25%	1%	5%
Average Duration[b]	< 5	10	< 10	10 +
Call Time	Off peak	Peak Time	business Hours	Any
Destination Call	Mostly local	Nation Wide	Local	International
Origin of Call	Home Cell	Any MSC/BSC	Same MSC	All
Type of Call	MTL[c]	MTL and **MTM**[d]	MTL	MTL and MTM

Table 1 : Classification of GSM subscribers

a. Value in minutes
b. Value in minutes
c. Mobile to Land
d. Mobile to Mobile

A simulation run can represent the behaviour of the network during one day. Therefore, a single Generator is able to simulate a population of 400 mobile stations during working days or week-ends.

4 Functionalities of IDAMN

The challenge was to model IDAMN in order to detect an intruder «***on-line***» while minimizing the overhead incurred in the cellular network. For this purpose, we have defined a set of analysis functions which are spread over the whole network. In fact, these functions are performed by different distributed entities of IDAMN in order to monitor mobile users.

The main novelty of our architecture is its ability to perform intrusion detection in the visited location and within the duration of a typical call as opposed to existing designs that require the reporting of all call data to the home location in order to perform the actual detection. In other words, the performance of IDAMN rely on the transfer of statistical information of small size instead of centralizing large amounts of audit data like in existing IDSs for mobile networks [16], [17].

4.1 Multi-Level Intrusion Detection

We have defined three different levels of analysis which can be summarized as follows:

- ***Level 1 - velocity and clone verification:*** this consists of a fast intrusion detection by verifying the speed of a mobile user or existing clones (the same user being active in two different parts of the network at the same time).
- ***Level 2 - component wise verification:*** the system measures the impact of the subscriber behaviour on the different GSM entities. For instance high activity of a switch in an area with a low density of inhabitants, may be a symptom of intrusion.

- ***Level 3 - intrusion detection per user***: this is the most significant intrusion detection analysis as the resulting procedure evaluates every deviation from the normal user's normal behaviour profile or «signature».

In the case of the level-3 analysis the basic idea of detecting an intruder relies on the system's ability to learn the normal behaviour of the subscriber by creating a ***user profile***. The profiles of IDAMN are designed to require a minimum amount of storage for historical data per mobile user and yet require relevant information that can efficiently be used during the detection process. An intruder impersonating the real subscriber will have a different behaviour and thus will generate a significant deviation from the standard profile [9].

In the case of GSM, the signature of the user is defined by 3 profiles: a mobility-profile, an activity-profile and a speech-profile[1]. Each profile will help in raising different intrusion alarms that a rule based system will analyse in order to give the final decision. Therefore, forbidding a mobile user to lend his mobile station becomes a mandatory security policy in this context.

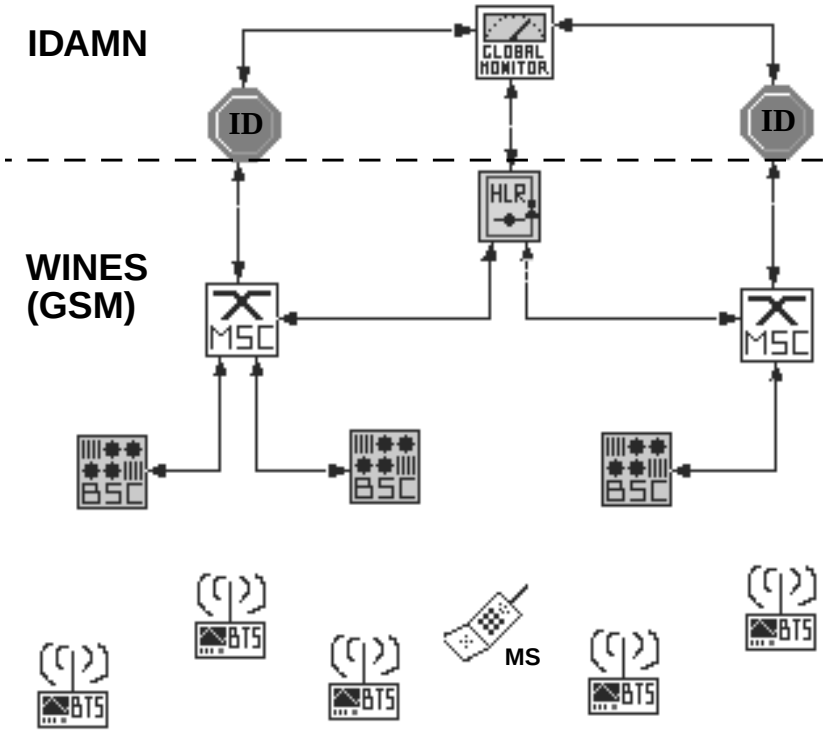

Figure 1 : Portion of the Global Platform

4.2 IDAMN Network Entities

The relevant information for modelling the user's behaviour as well as detecting an intrusion are distributed over the various entities of the GSM Network Sub-System. In order to minimize the computational overhead incurred in both IDAMN and GSM, the procedures addressing the different level of intrusion detection are spread over two dedicated machines: the ***Global Monitor*** (GM) and the ***Intrusion Detector*** (ID) as depicted in Figure 1.

1. *A speech verification algorithm is envisioned but has not been developed yet.*

The main purpose of the GM is to manage and to save over a long period the profiles of the subscribers. Moreover, the GM is in charge of executing the procedures corresponding to level 1 and level 2 of intrusion detection. The results of these 2 first analyses are processed by a rule based system which generates a *global report* to be passed to the ID.

The ID performs level 1 (detection of a clone at the location area level) and level 3 of intrusion detection. The ID is directly connected to the anchor[1] MSC and makes a copy of the relevant signalling messages generated by the mobile users. Depending on the nature of the messages, the ID initializes a set of statistical variables and requests the activity and mobility profiles from the GM as well as the global report concerning the mobile user. Then, the ID is able to detect a potential intrusion by comparing the variables to the received profiles according to some thresholds.

The GM is able to access the HLR via the GSM standard protocol MAP/C in contrast to the ID which can retrieve information from the VLR using the MAP/B protocol [24] if required. Moreover, a single ID is associated with each MSC allowing IDAMN to have a complete view on the mobile user in order to analyse his ***telephony activity*** and his ***migration pattern*** as discussed in the following sections.

5 Monitoring the Telephony Activity

The telephony activity of a mobile user is defined by two statistical vectors: the ***Call Vector*** and the ***Session Vector***. The values of these vectors are initialized by the ID upon receiving signalling messages from the anchor MSC as depicted in Figure 2.

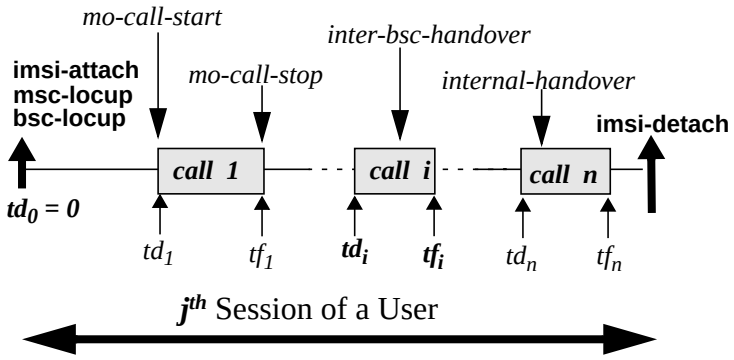

Figure 2 : Call Vectors initialized with GSM signalling messages

The Call Vector $\vec{V}_i^c$ is a local magnitude which models the i^{th} MO. We denote by td_i and f_i respectively the time when the MS triggered a *mo-call-start* and a *mo-call-stop* procedure. Upon receiving the GSM call management messages, the ID computes a statistical call vector over the duration of an outgoing call ($f_i - td_i$) , number of handovers performed[2] (nh_i) and inactivity between two calls ($d_i - tf_{i-1}$). The expression of $\vec{V}_i^c$ is as follows:

1. *In GSM terminology, this is an MSC managing the MS at the beginning of a call*
2. *Each time a handover procedure is performed, nh_i is increased by one*

$$\vec{}^{c}_{i} = \begin{bmatrix} td_i - tf_{i-1} \\ tf_i - td_i \\ nh_i \end{bmatrix} \quad (1)$$

The Session Vector $\vec{V}^s_j$ is a global magnitude which models the j^{th} session of the user in terms of network connection duration (D_j) total number of calls (Nc_j total duration of calls (Dc_j) and the total number of handovers Nh_j. The expression of $\vec{V}^s_j$ is the following:

$$\vec{}^{s}_{j} = \begin{bmatrix} D_j \\ Dc_j \\ Nh_j \\ Nc_j \end{bmatrix} \quad (2)$$

The high-speed intrusion detection algorithm which is able to interpret these vectors is based on the IDES approach [10]. However, our algorithm takes into account the specific context of wide cellular networks managing millions of subscribers such as GSM, whilst minimizing the computation overhead and the amount of data transferred between IDAMN entities.

In doing so, the telephony activity profile of a user is updated in a recursive way each time a new vector[1] $\vec{V}_n$ (i.e. $\vec{V}^s_n$ or $\vec{V}^c_n$) is initialized by the ID. Moreover, before incorporating a new vector $\vec{V}_{n+1}$ into the profile P_A, which is composed of a mean vector $\vec{M}_n$ and a covariance matrix C_n, we first compute $\vec{M}_n$ and C_n by multiplying each observed vector $\vec{V}_i$ by the following weight:

$$w_i = \frac{\alpha^{n-i}}{\sum_{i=1}^{n} \alpha^{n-i}} \quad (3)$$

where α is the forgetting factor ($0 < \alpha < 1$). This method creates a sliding time window for the observed vector which has the effect to favour the recent behaviour of the user. Therefore, if P_A has been previously computed over first *n* vectors (with *n* smaller than 30), the ID updates the mean vector as follows:

$$\vec{M}_{n+1} = \frac{1}{1-\alpha^{n+1}}(\alpha(1-\alpha^n)\vec{M}_n + (1-\alpha)\vec{V}_{n+1}) \quad (4)$$

and the covariance matrix is computed as follows:

$$\begin{aligned} C_{n+1} = {} & \frac{\alpha(1-\alpha^{n-1})}{1-\alpha^n} C_n + \frac{1-\alpha}{2}(\vec{M}_n - \vec{M}_{n+1})(\vec{M}_n - \vec{M}_{n+1})^t \\ & + \frac{1-\alpha^2}{2\alpha(1-\alpha^n)}(\vec{V}_{n+1} - \vec{M}_{n+1})(\vec{V}_{n+1} - \vec{M}_{n+1})^t \end{aligned} \quad (5)$$

1. *Whether it is a call vector or a session vector*

where *t* denotes the matrix transposition operation. Hence, the GM will not have to store all observed vectors in its database as the profile to be saved is $P_A = \left\{\overrightarrow{M_n}; C_n; \alpha\right\}$.

Once P_A has been generated, detecting an abnormal activity consists in computing the distance between a new vector $\vec{V}_{n+1}$ and the mean vector which is a function of the cluster of the *n* previous vectors. If this distance is smaller than a threshold S_{max} then, the new $\vec{V}_{n+1}$ is considered to be normal. Otherwise, the following inequation denotes the detection test for the abnormal case:

$$(\vec{V}_{n+1} - \vec{M}_n)^t C_n^{-1}((\vec{V}_{n+1} - \vec{M}_n) \geq S_{max}^2) \quad (6)$$

where C_n^{-1} represents the inverse matrix of C_n. We performed several experiments using different kinds of thresholds aiming to maximize the detection rate and to minimize the false alarm rate. Consequently, we obtain the best results with the following decision criterion:

$$S_{max} = max\left\{(\vec{V}_i - \overrightarrow{M_n})^t C_n^{-1}(\vec{V}_i - \overrightarrow{M_n})\right\}(for \quad i \in [1 \ldots n]) \quad (7)$$

The threshold S_{max} is not *a priori* fixed but is computed over the *n* first observed vectors for each user. Other thresholds have been used but the false alarm rate was high (between 20 to 50%) even if all intrusion scenarios were detected.

Category of user	Call false alarm rate	Call detection rate	Session false alarm rate	Session detection rate
DOMESTIC	1%	67 to 100%	2%	80 to 100%
BUSINESS	1%	88 to 100%	2%	90 to 100%
CORPORATE	1%	60 to 100%	5%	87 to 100%
ROAMER	2%	82 to 100%	1%	95 to 100%

Table 2 : False Alarm and Detection rates

For each mobile user belonging to a specific category of subscriber, 300 intrusive session vectors were tested (one session representing one day of connection to the network). Almost[1] every intrusive session was detected (at least 80%) with a false alarm rate varying from 2% to 5% as shown in Table 5. Much better performance was obtained with the call vector analysis (600 to 2000 vectors tested per user) as the false alarm rates were under 2% with a minimum detection rate of 60%. In every case, if a high intrusive activity is simulated, 100% of the matching vectors are detected; this corresponds to an intruder trying to misuse the network services extensively as already noticed in existing analog mobile networks.

6 Monitoring Users' Roaming

In addition to the telephony activity analysis, IDAMN includes a detection mechanism based on the monitoring of the users' travelling in the network. This mechanism It takes into account

1. *Depending on the category of user (domestic, corporate etc.)*

the areas visited by a subscriber as well as the most frequent itineraries followed. A mobility behaviour profile based on a graph model (with transition probabilities) is established depending on the frequency of location area crossings of the mobile user. This mobility profile is updated when the location updating procedures are performed (i.e, *inter-MSC-locup inter-BSC-locup*).

We assume that the migration of a mobile user is not completely random. For instance, most of the time a BUSINESS user will roam from his house to his office through a reasonable number of location areas and well defined cells. For the sake of simplicity, in the rest of the paper we denote by «cell» a location area.

Then, we associate to each cell a state, and compute the state transition probability. Depending on the itinerary[1] observed from the location update signalling messages, the probabilities of the different itineraries are calculated in order to give more importance to those the most followed. In other words, the transition probability going from a state *i* to a state *j* is a function of the number of times the mobile user has effectively traversed cell *i* to cell *j*.

Let X_n denote the stochastic variable indicating the cell on which the mobile station is located after the n^{th} location update. Consequently, if $X_n = 0$ the MS is considered to be disconnected and if $X_n = i$ the MS is camping on cell *i*. Therefore, given that at the previous ((n-1)'th) location update the mobile was camping on cell i, the transition probability from cell *i* to *j* is:

$$\begin{aligned} {}_{ij} &= p\langle X_n = j \mid X_{n-1} = i\rangle = \frac{n_{ij}}{n_{i_\circ}} \\ n_{i_\circ} &= \sum_k n_{ik} \end{aligned} \tag{8}$$

where n_{ij} is the number of times a user crossed cell *i* to cell *j* and $n_{i_\circ}$ is the total number of output occurrences of cell *i*. At the initial step, we consider that the mobile user can go in all directions with the same probability[2]. However, the training period will give higher probabilities to the most frequent itineraries depending on the migration habits of the user. Hence, the mobility profile of a user is a graph P_M where each cell is a state and each transition a probability as defined by Equation (8). In the example of Figure 3, the user crossed cells 2, 3 and 4 *n* times whereas cells 1, 3 and 5 were traversed only once.

The detection process consists of comparing the current itinerary in progress (called candidate itinerary) to the profile P_M in order to measure a strong deviation from that profile. Let $\hat{C}$ be the candidate itinerary composed of *k* cells; for each cell *i* of $\hat{C}$ we compute the average probability $\bar{p}_i$ as follows:

$$\bar{p}_i = \frac{1}{O_i} \cdot \sum_{\substack{j=1 \\ j \neq i}}^{n} p_{ij} = \frac{1 - p_i}{O_i} \qquad i \in [0 \ldots k] \tag{9}$$

where O_i is the total number of outgoing transitions of cell *i*.

1. *An itinerary is composed of successive cells or location area*
2. *All probabilities are initialized to zero*

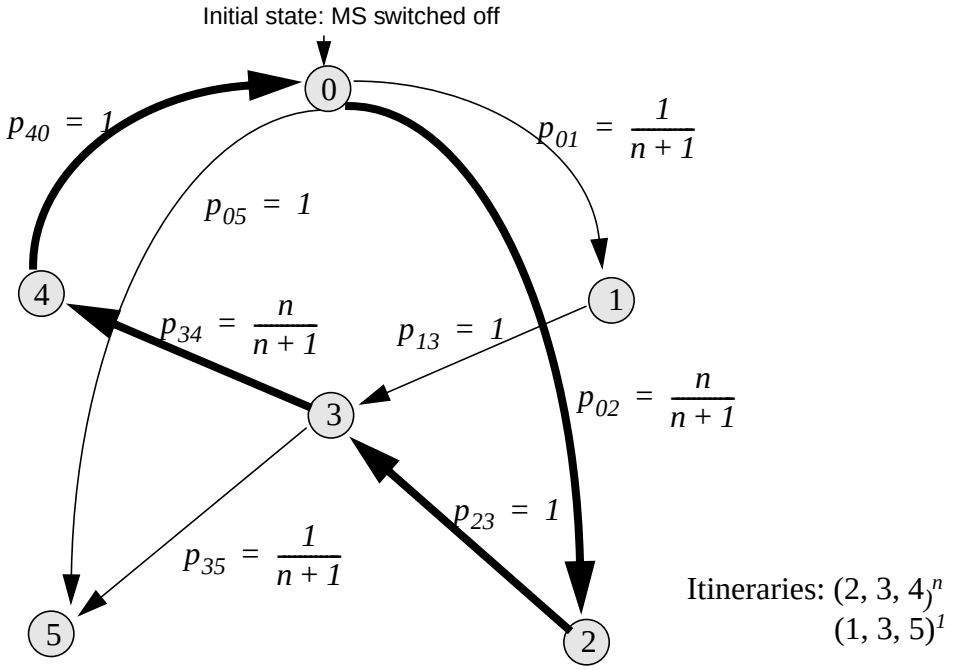

Figure 3 : Example of a Mobility Profile

The ultimate detection test is performed by analysing a doubt factor V_d ($V_d \geq 0$). Moving from cell i to cell j will be considered as abnormal if $p_{ij} \geq \bar{p}_i$. In that case, the doubt factor is reduced as follows: $_d = KV_d$ (with $0 < k < 1$). Otherwise, if $p_{ij} < \bar{p}_i$ then V_d is augmented with a positive value depending on the following situations:.

- If the cell has already been crossed then the doubt factor is increased by H as follows: $_d = V_d + H$.
- If the mobile unit arrives in a new cell for the first time V_d is incremented as follows: $V_d = V_d + \frac{F}{2^l}$,

where l represents the number of cells composing $\hat{C}$ which have never been encountered during earlier update of P_M.

As $\lim_{n \to \infty} \sum_{i=1}^{n} \frac{F}{2^i} = 2F$, we consider that the itinerary is abnormal if $V_d > 2F$. This decision criterion is tolerant to potential new itineraries of the legitimate user who may roam into visited domain instead of his usual home domain. This algorithm performs intrusion detection on-line as each time the mobile user performs a location update procedure it is possible to know whether or not the current cell traversed is normal.

In fact, parameters F and H depend on the average length of the user's itineraries. For each mobile user, 300 abnormal itineraries were tested over a population of 400 simulated users. The false alarm rate is lower than 15% with a detection rate varying from 65% to 100% as shown in Table 5.

Category of user	False alarm rate	Detection rate
DOMESTIC	1 to 2%	90 to 100%
BUSINESS	3 to 4%	75 to 90%
CORPORATE	4 to 6%	65 to 90%
ROAMER	5 to 7%	65 to 80%

Table 3 : False Alarm and Detection rates

7 IDAMN in Practice

Intrusion alarms are raised when the detector notices a strong deviation of the mobile user behaviour compared to the mobility and activity profiles. Next, these alarms are analysed by a rule based system which gives the final decision. If an intrusion is detected IDAMN can signal the network to activate a call bearing procedure or to disconnect the user.

The idea of comparing profiles of users with current session data is not exempt from potential problems. An intruder impersonating the subscriber can gradually modify his behaviour in order to escape detection. However, this implies that the attacker knows *a priori* the telephony activity as well as the migration pattern of the user. In a real large scale wireless network, it will be difficult for an attacker to obtain ***all*** audit data generated by the mobile user if a minimal service of privacy is provided by the means of encryption[1]. On the other hand, if the mobile network does not benefit from authentication and confidentiality mechanisms, then IDAMN must be provided with additional algorithms. For instance, intrusion detection algorithms based on speech verification or dialled number analysis will be an issue. Nevertheless, if we assume that an intruder having succeeded to obtain fraudulent access, his goal will be to use extensively the network services and thus will be automatically be detected considering the current version of IDAMN.

In order to demonstrate that implementing IDAMN in a real mobile network is feasible, we need to prove that the architecture is scalable and the algorithms have a reasonable complexity. Therefore, we denote by ***flop*** (floating point operation) an operation which require one multiplication and one addition: $1\ flop = S_{k-1} + a_k b_k$.

Thus, the number of flops needed to update the activity profile $P_A = \left\{ \overrightarrow{M_n}; C_n; \alpha \right\}$ is $(m^2 + 3m)/2$ where m is the dimension of the statistical vector. In comparison, without the recursive formula, the total number of flops needed to update P_A will be around $3m^2$. In the case of the intrusion detection test, we can also compute C_n^{-1} in a recursive way and hence, the number of flops needed to perform detection is $2m$.

With a DSP of the TMS 320 family [25], a floating point number is coded in 4 bytes and an integer in 2 bytes. Therefore, the global monitor of IDAMN will need $\left(\frac{m^2 + 3m}{2} + 1\right) \times 4$ bytes

1. *After being authenticated by the network the user shares a session key for the encryption of the signalling and user data.*

for a user's profile (P_A). The DSP TMS 320 is able to execute 33 million flops per second. In our case, we suppose that we need only 1 μs to execute 1 flop. Table 4 summarizes the characteristics of our algorithm.

Characteristics	Call Vector (m=3)	Session Vector (m=4)
Time to update a profile	9 µs	14 µs
Time for detection	6 µs	8 µs
Size of a profile	40 bytes	60 bytes

Table 4 : Characteristics of the activity analysis

If the profile P_M of a user is a graph composed of *m* states (or cells) then the complexity for creating or updating of P_M is m^2. The computation of $\bar{p}_i$ will need around *m* flops and the total number of operations to perform detection is $m^2 + m$.

As opposed to the telephony activity algorithm, the memory size to store P_M varies from a category of user to the other. For instance, a DOMESTIC user is rather static and the size of his mobility profile is smaller than the ROAMER user profile. Table 5 presents the characteristics of our algorithm.

Category of user	Number of cells[a]	Matrix size[b]	Detection Duration
DOMESTIC	4	32	16 µs
BUSINESS	10	200	100 µs
CORPORATE	16	512	256 µs
ROAMER	24	1152	576 µs

Table 5 : Characteristics of the migration analysis

a.In the graph
b.in bytes

The values shown in this table have been computed by considering the maximum number of cells traversed by each category of user. This algorithm might require more resources than the vector test algorithms but the time for the detection is also in the order of μs. Hence, the total size for both the activity and migration profile for one user is at most 1,5 Kb.

A «***real world***» implementation of IDAMN can be envisioned. The planning of IDAMN depends on the topology of the mobile network and in our case on GSM. In general, a HLR manages 500 000 subscribers and a MSC 40 000 subscribers [1]. Therefore, if the network comprises 1 000 000 mobile users[1], IDAMN will be composed of 2 GMs and 25 IDs which are interconnected with a 64 Kbits/s leased line and each GM managing at most 13 IDs.

1. *For instance, 800 000 subscribers are currently using the GSM network of the french operator France Telecom*

According to [1], during peak hours a MSC receives in average 24 MO calls per second but for demonstration, we consider the worst case of 40 MO calls per second. If we assume that this load is the same for the 13 MSCs, then each ID sends in 65 ms $40 \times 100 = 4$ Kbits of profile requests[1] to one GM. Upon receiving these profile requests, the GM must search for the 40 profiles (1 ms needed per user profile[2]) in order to transmit them to the corresponding ID. The time needed to perform these operations is around $1,5 \times 40 / 64 \approx 0,9\ s$. Hence, because an ID takes at most 600 μs to detect an abnormal behaviour, all intrusion detection algorithms can be performed in less than 1 second.

Note that the user profiles are transmitted only once between the GM and an ID. Upon receiving new activity or mobility data, the same ID will be able to perform detection immediately without contacting the GM. Moreover, IDAMN does not save any call detail records because if the call is normal then the ID updates the profile and the call data is dropped. This has the advantage to minimize the amount of storage compared to centralized IDSs which have to save several days of call detail records (cellular billing information) in order to allow the fraud investigator to view the call activity of the mobile user.

8 Conclusion

This paper presented IDAMN a distributed intrusion detection system for cellular networks. IDAMN is able to have a complete overview on the mobile user telephony activity and mobility behaviour. The main novelty of our architecture is its ability to transfer a small user profile near the location of the serving MSC as opposed to existing systems which require the reporting of large amount of audit data to a central monitoring entity. The algorithms have been designed in order to create and update the user profiles in real time without having *a priori* knowledge of the behaviour of the user. Moreover, an alarm can be raised within the duration of a typical call. Preliminary testing of the components of IDAMN are promising as we can raise an alarm in less than one second with an intrusion detection rate varying from 70% to 100% whereas the false alarm rate is lower than 5%. Future work includes and enhancement of IDAMN's functionalities by developing speech and dialled number verification algorithms.

9 Acknowledgments

We are indebted to Jacques Labetoulle and Dirk Slock for their insightful comments on previous versions of the intrusion detection algorithms. We would like also to thank Raymond Knopp for his assistance during the preparation of this article and Christian Bonnet for the implementation of the GSM simulator WINES.

References

[1] M. Mouly, M.B. Pautet, «The GSM System for Mobile Communications», ISBN 2-9507190-0-7, 1993.

[2] R. Molva, D. Samfat and G. Tsudik, «Authentication of Mobile Users», IEEE Network Magazine, Special Issue on Mobile Communications, March/April 1994.

[3] Ashar Aziz and Whitfield Diffie, «Privacy and Authentication for Wireless Local Area Networks», IEEE Personal Communication, First Quarter 94.

[4] M. Beller, L. Cheng and Y. Yacobi, «Privacy and Authentication on a Portable Communication System»,

1. *In the protocol we have implemented, the size of the profile request is 100 bits.*
2. *This is a time needed to a VLR to retrieve information in the database.*

IEEE Journal on Selected Area in Communication, Aug. 1993.

[5] B. Mukherjee, L.T. Herbelein, K.N. Levitt, Network Intrusion Detection, IEEE Network Magazine, May/ June 1994, VOL. 8 No 3.

[6] G.B. White, E.A. Fish and U.W. White, «Cooperating Security Managers: A Peer-Based intrusion Detection System», IEEE Network Magazine, January/February 1996.

[7] D. Samfat, V. Devernay, C. Bonnet, «A GSM Simulation Platform for Intrusion Detection», Proceedings of ICC'95, Seattle, June 1995.

[8] J. Hochberg et al., « NADIR: an Automated System for Detecting Network Intrusion and Missue», Computer and Security, vol. 12, no 3, May 1993.

[9] T.F. Lunt, "IDES: An Intelligent System for Detecting Intruders", Proceedings of the Symposium: Computer Security, Threat and Countermeasures, Rome, Italy, November 1990.

[10] Theresa F. Lunt, «Using Statistics to Track Intruders», Proceedings of the Joint Statistical Meetings of the American Statistical Association, August 1990.

[11] T.F. Lunt, «A Real-time Intrusion Detection Expert System (IDES)», Interim Progress Report, Project 6784, SRI International, May 1990.

[12] K. Ilgun, P.A Poras, R.A Kemmerer, «USTAT a Real-Time Intrusion Detection System For Unix», SRI Intrusion Detection Workshop, Paolo Alto, CA Feb. 1989

[13] S.E. Smaha, «Haystack: An Intrusion Detection System», Proceedings of the IEEE Fourth Aerospace Computer Security Applications Conference, Orlando, FL, Dec. 1988

[14] Steven R. Snapp et al. , "DIDS (Distributed Intrusion Detection System) - Motivation, Architecture and an Early Prototype", Proceedings of the 14th National Computer Security Conference, Washington, DC, Oct 91.

[15] L.T. Hebberlein, K.N Levitt and B. Mukherjee, "An Intrusion-Detection System for Large-Scale Networks", Proceedings of the 15th National Computer Security Conference, Baltimore, MD, October 1992.

[16] Rob Mechaley and Kirk Calson, «Parameters for Fraud Management Using Network Based Techniques», Technical Report 45.2.II.2, Mc Caw Cellular Communications, Inc, 1991.

[17] Evan J. Davies, Henrik Nordin, Sharon Hershon, Kathy Gallup, «Design Overview for Fraud Detection and Analysis System», Technical Report, Digital Equipment Corporation, August 1992.

[18] ETS 300 175-7, «Part 7: Security Features», DECT common Interface, ETSI, Oct. 1992.

[19] «Cellular Digital Packet Data (CDPD)», System Specifications, Release 1.0, July 1993

[20] GSM Recommandation 3.20, «Security Related Network Function», ETSI Standard, Feb. 1992.

[21] GSM Recommendation 3.05, Signalling Requirements Relating to Routing of calls to Mobile Subscribers, ETSI Standard, Feb. 1992.

[22] GSM Recommendation 3.09, Handover Procedures, ETSI Standard, Feb. 1992.

[23] GSM Recommendation 3.12, Location Registration Procedures, ETSI Standard, Feb. 1992.

[24] GSM Recommendation 9.02, MAP Specification, ETSI Standard, Feb. 199.

[25] Panos Papamichalis, "Digital Signal Processing Application with the TMS320 Family", Application Book, Texas Instruments 1990.

[26] OPNET Network Simulation Software, MIL 3, Inc. 3400 International Drive, Washigton, DC 20008, Release 2.4.A.