Project Number	AC095
Project Title	ASPeCT: Advanced Security for Personal Communications Technologies
Deliverable Type	M (Major)

Deliverable Number	AC095/KUL/W22/DS/P/06/A
Contractual Date of Delivery	August 1996 (Y2M6)
Actual Date of Delivery	August 1996 (Y2M6)
Title of Deliverable	Definition of Fraud Detection Concepts
Contributing Work Packages	WP2.2
Nature of the Deliverable	R (Report)
Editors	Yves Moreau and Bart Preneel, K.U.Leuven

Abstract	This report describes the different techniques developed within the ASPeCT project to detect fraud in a mobile communication network. This document discusses the following topics: • Description of fraud patterns to be detected using advanced techniques • Description of network data that will be used and its pre-processing (including sanitisation) • User profiling: approach and implementation • Rule-based approach to fraud detection • Neural network based approach to fraud detection: unsupervised learning • Neural network based approach to fraud detection: supervised learning • Legal aspects of fraud detection The approaches described in this document will be implemented in a set of first prototypes, which will be described in Deliverable 08 - together with a performance assessment of the different techniques.

Keywords	ACTS, ASPeCT, fraud detection

0. Executive Summary

It is estimated that the mobile communication industry loses several million ECUs per year due to fraud. Therefore, prevention and early detection of fraudulent activity is an important goal for network operators. It is clear that the additional security measures taken in GSM and in the future UMTS (Universal Mobile Telecommunications System) make these networks less vulnerable to fraud than the analogue networks. Nevertheless, certain types of commercial fraud are very hard to preclude by technical means. The use of sophisticated fraud detection techniques can assist in early detection of such frauds, and will also reduce the effectiveness of technical frauds.

A first step of the work consists of the identification of possible fraud scenarios in telecommunication networks and particularly in mobile networks. In [3], these scenarios have been classified by the technical manner in which they are committed; also an investigation has been undertaken to identify which parts of the mobile telecommunication network are abused in order to commit any particular fraud. Subsequently, typical indicators that may be used to detect fraud committed using mobile telephones have been identified. In order to provide an indication of the likely ability of particular indicators to identify a specific fraud, these indicators have been classified both by their type and by their use.

The different types are as follows:

- Usage indicators, related to the way in which a mobile telephone is used
- Mobility indicators, related to the mobility of the telephone
- Deductive indicators, which arise as a by-product of fraudulent behaviour (e.g., overlapping calls and velocity checks)

The different uses are the following:

- Primary indicators can, in principle, be employed in isolation to detect fraud.
- Secondary indicators provide useful information in isolation (but are not sufficient by themselves).
- Tertiary indicators provide supporting information when combined with other indicators.

A selection has been made of those scenarios which cannot be easily detected using existing tools, but which could be identified using more sophisticated approaches.

The potential fraud indicators have been mapped to network data required to measure them. A great deal of the information required to monitor the use of the communications network is contained in the toll tickets (supplementary information is contained within the associated signalling information). A toll ticket is a set of details stored electronically

about any individual call being made on a mobile telephone. It is used to determine the charge to the subscriber, but it also provides information about customer usage and thus facilitates the detection of any possible fraudulent use. It has been investigated which fields in the GSM toll tickets can be used as indicators for fraudulent behaviour. In order to protect the privacy of users during the development of the fraud detection techniques within ASPeCT, all personal information in the toll tickets (such as telephone numbers) is encrypted.

More sophisticated frauds do not become apparent from a single toll ticket. In order to detect such frauds, one needs to analyse both absolute uses of the network as well as changes in user behaviour. The information on the behaviour of a user is monitored over two windows: the short term window is called the Current User Profile (CUP), while the long term window is called the User Profile History (UPH). These profiles contain condensed information rather than the full sequence of toll tickets. When a new toll ticket arrives, both profiles are updated, in such a way that information on previous toll tickets is being decayed.

In ASPeCT, several approaches are taken to identify fraudulent behaviour. In the rule-based approach, both the absolute and differential usage are verified against certain rules. This approach works best with user profiles containing explicit information, where fraud criteria given as rules can be referred to. User profiles are maintained for the directory number of the calling party (A-number), for the directory number of the called party (B-number) and also for the cells used to make/receive the calls. A-number profiles represent user behaviour and are useful for the detection of most types of fraud, while B-number profiles point to hot destinations and thus allow the detection of frauds based upon call forwarding. All deviations from normal user behaviour identified via the different analysing processes are collected and alarms will finally be raised if the results in combination fulfil given alarm criteria. The implementation of this solution is based on an existing rule-based tool for audit trail analysis.

A second approach to identify fraudulent behaviour uses neural networks. An important element here is the reduction of the dimensions of the behaviour space. Both linear techniques (such as principle components analysis) and non-linear dimension reduction based on auto-associative feed-forward neural networks are being developed. As a consequence, the parameters in the profile no longer have a simple interpretation. Subsequently, the users are classified based on their usage in order to enhance the sensitivity of the system; tariff class forms a first element, but a more sophisticated classification based on the self-organising properties of neural networks is being developed. Finally, fraudulent behaviour is identified by supervised learning of a neural network.

1. CONTENTS

2. Document Control

2.1 Document History

This is the final draft of Deliverable 06.

2.2 Change Control

In conformance with the project Quality Plan [1].

3. Abbreviations and Glossary of Terms

ASPeCT	Advanced Security for Personal Communications Technologies
AU	Analysing Unit
AUA	Absolute Usage Analyser
CU	Controlling Unit
CUP	Current User Profile
DUA	Differential Usage Analyser
EIR	Equipment Identification Register
GUI	Graphical User Interface
GSM	Global System for Mobile Communications
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
MA	Master Analyser
MSC	Mobile Services Switching Centre
MSISDN	Mobile Station Integrated Services Digital Network
PABX	Private Automatic Branch Exchange
PACE	Police and Criminal Evidence Act
PDAL	Protocol Data Analysis Language
PDAT	Protocol Data Analysis Tool
PSTN	Public Switching Telephone Network
RCF	Roaming Call Forward
SIM	Subscriber Identity Module
SP	Service Provider
TACS	Total Access Communications System
TMN	Telecommunications Management Network
TT	Toll Ticket
UMTS	Universal Mobile Telecommunications System
UPR	User Profile Record
UPH	User Profile History
WORM	Write Once - Read Many

4. Introduction

This document is organised as follows:

- Description of fraud patterns to be detected using advanced techniques
- Description of network data that will be used and its pre-processing (including sanitisation)
- User profiling: approach and implementation
- Rule based approach to fraud detection
- Neural network based approach to fraud detection: unsupervised learning
- Neural network based approach to fraud detection: supervised learning
- Legal aspects of fraud detection

The approaches described in this document will be implemented in a first set of prototypes; these prototypes will be described in Deliverable 08, together with a performance assessment of the different techniques.

5. Description of Fraud Patterns to be Detected using Advanced Techniques

In this section, we give an overview of the most common types of fraud patterns, together with an outline of their characteristics as seen by the Network Operator. This section has been withheld from any public dissemination because of commercial confidentiality considerations. Accordingly, it has been classified as ASPeCT Confidential. The contents are moved to Annex A.

For each type of fraud identified, the fields of the toll tickets that maybe helpful for fraud detection are given. The description of the toll ticket fields conform to the Vodafone Archived Eurobill format.

The various types of fraud are classified according to whether they are technical fraud operated for financial gain, or they are fraud related to personal use - hence not employed for profiteering. A further classification is achieved by considering whether the network abuse is the result of administrative fraud, procurement fraud or application fraud.

6. Description of Network Data and its Pre-processing (including sanitisation).

In this section, we describe the data produced by the network, and how this data is pre-processed. This includes a description of how the data is sanitised in order to preserve the privacy of the users. This section has been withheld from any public dissemination because of commercial confidentiality considerations. Accordingly, it has been classified as ASPeCT Confidential. The contents are moved to Annex B.

This section includes an overview of the requirements of the ASPeCT partners developing the fraud engines and of the Network Operators / Service Providers. A recommended field format for the sanitisation of the toll tickets is described in detail, together with the rationale for the selection of fields to be sanitised and their respective lengths to be sanitised.

This is followed by a detailed description of the respective techniques for sanitising the toll tickets, for padding toll ticket fields to a standard length, for extracting the relevant data from the sanitised toll ticket fields and for specifying toll tickets for special retrieval. The latter requirement allows encrypted toll ticket fields, associated with a particular (anonymous) subscriber, to be used to specify (to the respective Network Operator) the user for whom further retrieval of (archived) toll tickets is desired by the fraud engine development partners.

This section is concluded with a summary of the special toll ticket format used by the ASPeCT partners, together with an example of the input and corresponding output of the toll ticket sanitisation program.

7. User Profiling: Approach and Implementation

7.1 Absolute or Differential Analysis

Toll Tickets are data records containing details pertaining to every mobile phone call attempt that is made. Toll Tickets are transmitted to the network operator by the cells or switches that the mobile phone was communicating with at the time due to proximity.

In addition to providing necessary billing information, Toll Tickets contain a wealth of information that can be used to catch a fraudster. Chapter 6 introduced a number of fraud scenarios that have been encountered in analogue networks and detailed the characteristics of Toll Ticket parameters observed whilst the fraud is being committed. Existing fraud detection systems tend to interrogate sequences of Toll Tickets comparing a function of the various fields with fixed criteria known as *triggers*. A trigger, if activated, raises an alert status which cumulatively would lead to an investigation by the network operator. Such fixed trigger systems perform what is known as an *absolute* analysis of the Toll Tickets and are good at detecting the extremes of fraudulent activity.

Another approach to the problem is to perform a *differential* analysis. Here we monitor behavioural patterns of the mobile phone comparing its most recent activities with a history of its usage. Criteria can then be derived to use as triggers that are activated when usage patterns of the mobile phone change significantly over a short period of time. A change in the behaviour pattern of a mobile phone is a common characteristic in nearly all fraud scenarios excluding those committed on subscription where there is no behavioural pattern established.

There are many advantages to performing a differential analysis through profiling the behaviour of a user. Firstly, certain behavioural patterns may be considered anomalous for one type of user, and hence potentially indicative of fraud, that are considered acceptable for another. With a differential analysis flexible criteria can be developed that detect any change in usage based on a detailed history profile of user behaviour. This takes fraud detection down to the personal level comparing like with like enabling detection of less obvious frauds that may only be noticed at the personal usage level. An absolute usage system would not detect fraud at this level. In addition, however, because a typical user is not a fraudster, the majority of criteria that would have triggered an alarm in an absolute usage system will be seen as a large change in behaviour in a differential usage system. In this way a differential analysis can be seen as incorporating the absolute approach.

7.2 The differential approach

Most fraud indicators do not become apparent from an individual Toll Ticket. With the possible exception of a velocity trap, we can only gain confidence in detecting a real fraud through investigating a fairly long sequence of Toll Tickets. This is particularly the case

when considering more subtle changes in a user's behaviour by performing a differential analysis.

A differential usage system requires information concerning the users history of behaviour plus a more recent sample of the mobile phones activities. An initial approach might be to extract and encode information from Toll Tickets and to store it in record format. This would require two windows or spans over the sequence of transactions for each user. The shorter sequence might be called the Current User Profile (CUP) and the longer sequence, the User Profile History (UPH). Both profiles could be treated and maintained as finite length queues. When a new Toll Ticket arrives for a given user, the oldest entry from the UPH would be discarded and the oldest entry from the CUP would move to the back of the UPH queue. The new record encoded from the incoming Toll Ticket would then join the back of the CUP queue.

Clearly it is not optimal to search and retrieve historical information concerning a user's activities prior to each calculation, on receipt of a new Toll Ticket. A more suitable approach is to compute a single cumulative CUP and UPH, for each user, from incoming Toll Tickets which can be stored as individual records, possibly in a database. So that we maintain the concept of having two different spans over the Toll Tickets without retaining a database record for each Toll Ticket, we will need to decay both profiles before the influence of a new Toll Ticket can be taken into consideration. A straight forward decay factor may not be suitable as this will potentially dilute information relating to encoded parameters stored in the user's profile. Watering down parameters such as call duration or distances, if performing a geographical analysis of B numbers, might result in a user's profile exhibiting false behaviour patterns. Each of the three systems, to be detailed in the following sections, will handle the decaying, or fading, technique in slightly different ways.

7.3 Relevant Toll Ticket Data

There are two important requirements for user profiling. At first, efficiency is of the foremost concern for storing the user data and for performing updates. Secondly, user profiles have to realise a precise description of user behaviour to facilitate reliable fraud detection. All the information that a fraud detection tool will need to handle is derived from the toll tickets provided by the network operator.

The following toll ticket components have been viewed to be the most fraud relevant measures:

- Charged_IMSI (identifies the user)
- First_Cell_Id (location characteristic for mobile originating calls)
- Chargeable_Duration (base for all cost estimations)
- B_Type_of_Number (for distinguishing between national / international calls)
- Non_Charged_Party (the number dialled)

These components will continually be picked out of the toll tickets and incorporated, in some way, into the user profiles in a cumulative manner.

7.3.1 Building features from Toll Ticket parameters.

Using the Toll Ticket parameters described in the previous section, we can derive useful features that would be suited to an analysis of user behaviour profiles. The most promising features that could be encoded are:

- IMSI
- Number of originated national Calls per Time Interval
- Number of originated international Calls per Time Interval
- Total Duration of originated national Calls per Time Interval
- Total Duration of originated international Calls per Time Interval
- Number of hot Destinations per Time Interval
- Statistical Measures (variance, ..) per Time Interval
- Fraud Ranking (computed from number and severity of past events)

With this specification we have succeeded in using only one type of data, namely numbers. This hopefully will make updating the computed user profiles a lot easier.

Categorical information like the B_Type_of_Number leads to doubling some components like the number of calls and the total duration. Hot destinations will initially be provided by the network operator and later such lists will be enlarged by the fraud detection tool itself.

7.3.2 Suitable parameters for a B number analysis.

Information concerning the destination of a call, also stored in the Toll Ticket, will facilitate the detection of fraud scenarios with a high frequency of calls to the same B_numbers. In this case, the B-number can be either a fraudster using supplementary services (call forwarding, conference calls,...) or a victim of fraud (e.g. a PABX or a free-phone service). The following measures are considered important to a B number analysis:

- Non_Charged_Party (MSISDN of B-party)
- Number of incoming Calls per Time Interval
- Total Duration of incoming Calls per Time Interval
- Number of hot Origins per Time Interval
- Statistical Measures (variance, ..) per Time Interval
- Fraud Ranking (computed from number and severity of past events)

Potentially, destination profiles accumulating evidence of the above parameters could be constructed when the corresponding supplementary services are used. They would be applied not only for mobile users and free-phone numbers of the mobile network but also for known PABXs within the PSTN.

7.3.3 Potential for the analysis of cell congestion.

Analysing cell congestion will usually not lead to statements of the same value as those derived from A and B-number analysis. However, they can provide ancillary information in connection with further analysis. Geographical information might also be useful for the detection of the fraudsters themselves.

8. Rule Based Approach to Fraud Detection

The rule based fraud detection tool is based on a tool called PDAT. PDAT (Protocol Data Analysis Tool) is a rule based tool for intrusion detection, which has been developed by Siemens ZFE (Corporate Research and Development). PDAT works in heterogeneous environments, has the possibility of on-line analysis, and provides a performance of about 200 KB input per second. While the actual analysis is performed by one or more analysing units (AU), the control of analysis is centralised in a Controlling Unit (CU). Important goals were flexibility and broad applicability, including the analysis of general protocol data, which is achieved by the special language PDAL (Protocol Data Analysis Language). PDAL allows the programming of analysis criteria as well as a GUI-aided configuration of the analysis at runtime.

Intrusion detection and mobile fraud detection are quite similar problem fields and the flexibility and broad applicability of PDAT are promising for using this tool for mobile fraud detection too. The main difference between intrusion detection and mobile fraud detection seems to be the kind of input data. The recording for intrusion detection produces 50 MB per day and user, but only for the few users of one UNIX-system. In comparison, fraud detection has to deal with a huge amount of mobile phone subscribers (roughly 1 Million), each of whom, however, produces only about 300 bytes of data per day. PDAT was able to keep all interim results in main memory, since only a few users had to be dealt with. For fraud detection, however, intermediate data has of course to be stored on hard disc. Because of these new requirements it was necessary to develop some completely new concepts. One of these new concepts is the user profiling introduced in the next section. Another concept is the fast swapping method described in Section 8.3. Finally, the internal architecture had to be changed to a great extend. The next sections should show, that these are effective measures of migration and enlargement and will provide us with a powerful fraud detection tool.

8.1.1 Mapping of User Behaviour to Profiles

Even over a long period of time, user behaviour should be storable efficiently without loss of essential information. For this purpose we use for the short-term past and for the long-term past two different methods of user profiling. While user information gathered in the recent history is represented by current user profiles (CUP), information collected in the long term is stored in the user profile history (UPH). For the short-term past several CUPs can be stored to keep more detailed information for a longer time period. The period can be split into *m* time intervals, for each of which we will build a CUP.

This results in a sequence of CUPs $(CUP)_i$, $i = 1,..,m$, with CUP_i being the profile of the i-th last time interval. CUP_1 is the profile of the current time interval. Thus it is the newest profile of all and is usually not finished. (A reasonable choice for the time interval length could be one day.) This way we gain a window for a time period of adjustable length (m). The window provides more information than single user profile does. It also tells us the

course (a story) of the user behaviour during a time period. This enables further statistical computations such as, for example, estimating the means, variances, covariances of the different quantities used in the user profile. Besides, this method allows more thorough analysis for new subscribers.

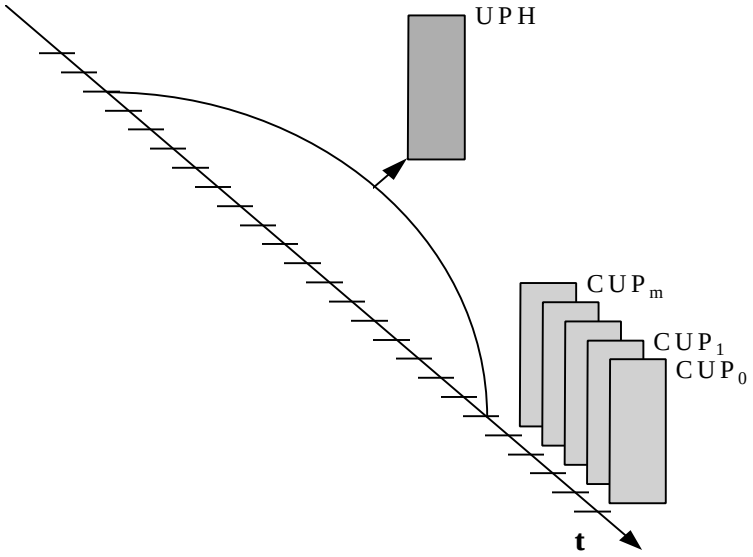


Figure 8-1 Mapping the User Behaviour to User Profiles during a Period of Time

8.1.2 Fading

In the interests of memory efficiency, the sequence of CUPs is not expanded into the long-term past. For the long-term past we maintain one single user profile, the user profile history (UPH). Since user behaviour might considerably change in a longer time, the UPH must be changeable, too. This is done by decaying the current UPH-values and adding "fresh" data taken out of one or more CUPs. This process is called fading. The following equation describes exponential fading in general:

$$UPH_{new} = (1\text{-}f)\ UPH_{old} + f \sum_{i=1,..,m} w_i\ CUP_i \text{, with fading-factor } f \in \mathbf{R},\ 0 \le f < 1$$

and weights $w_i \in \hat{\mathbf{A}}$, $0 \le w_i \le 1$ for $i = 1,.., m$.

The multiplication of profiles by a certain factor is meant as a multiplication of all components of this profile. The term $(1\text{-}f)\ UPH_{old}$ shows how fading is applied to the old *UPH*: A low fading factor e.g. $f = 0.01$ will decay the old values by 1 percent at each update. Additionally 1 percent of new CUP information is added to the new UPH. The factor f and the weights w_i should be chosen in a way that allows the differential analysis to stay sensible enough for still detecting unusual deviations of user behaviour.

Two examples are given below:

1. Exponential fading with oldest CUP: ($w_i = 0$, for $i = 1,..,m\text{-}1$ and $w_m = 1$)

$$UPH_{new} = (1\text{-}f)\ UPH_{old} + f\ CUP_m, \quad \text{with fading-factor } f \in \hat{\mathbf{A}},\ 0 \le f < 1$$

2. Exponential fading with CUP-Average

$$UPH_{new} = (1\text{-}f)\ UPH_{old} + f \sum_{i=0,..,m-1} CUP_i/m, \text{ with fading-factor } f \in \hat{\mathbf{A}},\ 0 \le f < 1$$

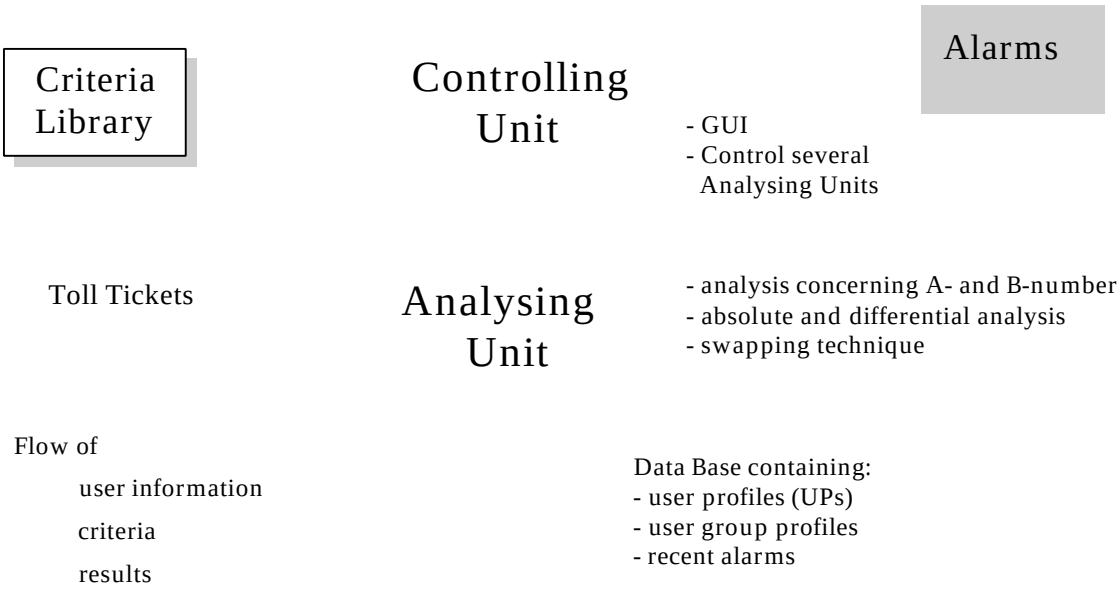
A characteristic of exponential fading is that the decay is exponential relating to *n*, where *n* is the number of iterations:

$$UPH_n = (1\text{-}f)^n\ UPH_0 + (1\text{-}f)^{n-1}\ f\ CUP_{m1} + ...+ f\ CUP_{mn}$$

in the case of Example 1.

8.2 Architecture and Functionality of the Rule Based Fraud Detection Tool

The general architecture of the rule based fraud detection tool is shown in the figure below. There are two active units: the Controlling Unit (CU) and the Analysing Unit (AU). The Controlling Unit provides a GUI, which is the sole interface to the tool's administrator. The CU has access to a Criteria Library where all rules used for operation of the tool are stored. The criteria consist mainly of two categories, the first one comprises rules needed for operational purposes like profile updating, while the second category contains the actual criteria for fraud detection. When starting the fraud detection tool, the criteria library is read from the CU and transferred to the AU. Additionally, during operation new rules can be added to the library via the CU and will immediately take effect. The second job of the CU is reporting all fraud alarms. In general the CU can control several AUs which all can be located on different machines. However, this feature will not be needed for our tool. The analysing unit is the essential part of the fraud detection tool. Here all toll tickets are processed and all alarms will be detected as well. The AU also performs the user profiling concerning A-number, B-number and cells (1st prototype: A-number analysis only) using a fast swapping technique in order to efficiently store profiles to a data base. Based on the profiles the AU allows absolute and differential analysis as well.


Criteria Library
Controlling Unit
- GUI
- Control several Analysing Units
Alarms
Toll Tickets
Analysing Unit
- analysis concerning A- and B-number
- absolute and differential analysis
- swapping technique
Flow of
user information
criteria
results
Data Base containing:
- user profiles (UPs)
- user group profiles
- recent alarms

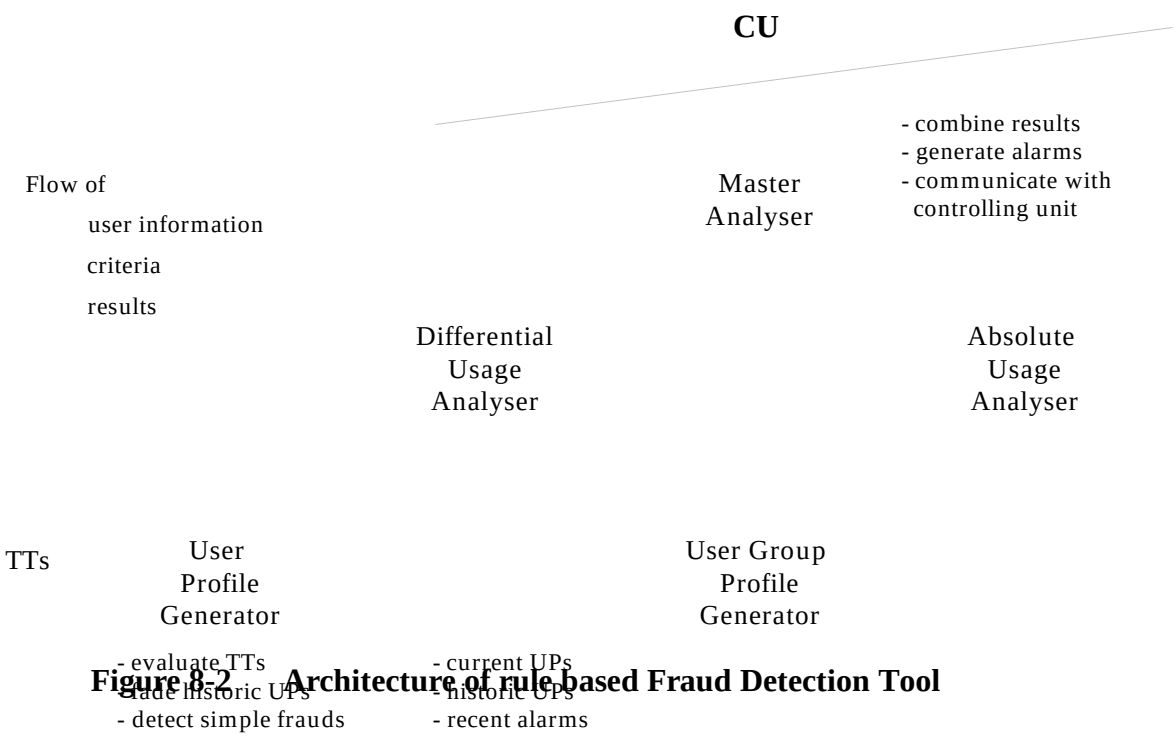


Figure 8-2 Architecture of rule based Fraud Detection Tool

A more detailed view of the AU is shown in Figure 9-2. The AU is divided up into 5 functional parts. The User Profile Generator is the front-end of the tool. Here incoming toll tickets are queued and finally adapted to an internal format in order to extract the essential information and to allow references. The swapping technique is used to keep toll ticket information to a specific amount in main memory and allows updates of bigger portions to the user profiles on hard disc. Periodically, fading is issued by the front-end to update the UPHs, too. The Differential Usage Analyser (DUA) works on both the CUPs and the UPH of a user. This way changes in the user's behaviour can easily be detected. The Absolute Usage Analyser (AUA) compares the current behaviour with some absolute usage thresholds. Absolute usage thresholds can also be refined by using group specific thresholds. This method and the necessary user group profile generator, however, will not be implemented for the first prototype.

Beyond these analysing parts there is a so-called Master Analyser (MA), which firstly combines all the results and decides when to generate an alarm. The second task of the MA is the communication with the CU. In the one direction the MA transmits generated alarms to the CU; in the other direction it receives all rules from the CU and distributes them to the other functional parts.

9.5 Neural Network Architecture of the Fraud Detection Engine

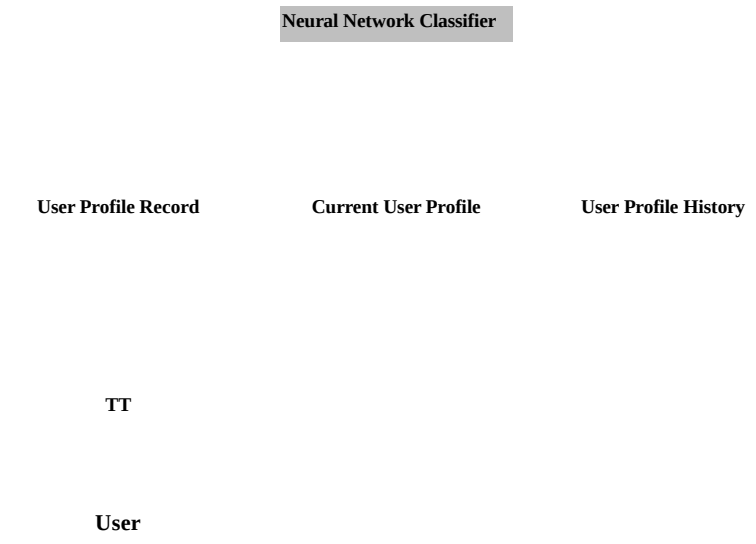
The fraud detection engine is based on the following architecture. A user is uniquely identified by its IMSI. To each user is associated a Current User Profile (CUP) and a User Profile History (UPH). The Current User Profile contains the short-term behaviour of the user while the User Profile History contains the long-term behaviour of the user. We present the architecture in two parts: the architecture used for training the neural network and the architecture used for run-time fraud detection.

9.5.1 Training architecture

Let us first consider the case of a single user. When a new toll ticket is provided, a feature vector called the User Profile Record (UPR) is extracted from it. This is done by extracting the relevant fields of the toll ticket. For the categorical variables that have too many categories, they can be grouped into larger categories (for example, foreign countries can be grouped by continent). Then, the categorical variables are encoded, and the ordered variables are standardised. The User Profile Record obtained in this way is used to update the Current User Profile. The reason for using CUPs and UPHs instead of keeping the whole list of calls is that it would impose totally unrealistic memory requirements. The update can thus be done according to two strategies (depending on which feature needs to be updated). In the first case, the new value in the CUP is a weighted sum of the old value and of the corresponding value of the UPR (e.g., $CUP(i+1) = 0.95*CUP(i) + 0.05*UPR$). Great care might be needed in order to find the appropriate value for the weighting factor. Too small a contribution from the UPR and the CUP will never be able to adapt itself, too large a contribution from the UPR and the CUP will not be able to memorise the behaviour of the user. Also, there are interesting quantities that cannot be easily handled this way. For example, one might want to know the total duration of the calls in the last twenty-four hours. Using the first strategy, one could compute a short-term estimate of the average call duration and a short-term estimate of the average number of calls per day. The product of the two would be an estimate of the total call duration in the last twenty-four hours. However, this is a cumbersome approach. Another strategy would be to form blocks by adding up the call durations during twelve hour periods (12am-12pm, 12pm-12am) and, let us say, retain only the last four blocks. Once one goes past twelve o'clock, the oldest block is thrown away and a new block is begun. Similarly, one updates the User Profile History using the Current User Profile, although this does not necessarily have to be done with each UPR, but can be delayed for a number of toll tickets or for some period of time.

User profiling is not the only profiling option. Although this will not be implemented in the first prototype, another important profiling technique is B-number profiling. When a toll ticket is presented to the system, the corresponding user profiles are loaded in the fraud engine. But we could also maintain profiles for B-numbers belonging to sensitive destinations or for recently called B-numbers (a database with all B-numbers would be too large). Those profiles would also be loaded in the fraud engine. This would enhance the performances for the detection of frauds, such as PABX fraud.

Training of the system is achieved as described in the paragraph on pattern recognition theory. The data is presented user by user. For each user, all its toll tickets are presented one by one to the network. The targets are defined as explained in the paragraph on data labelling. The weights are then adapted using a gradient descent algorithm. One pass consists in presenting all the available data (minus the validation and test data). Passes are done until the error on the validation set reaches a minimum.



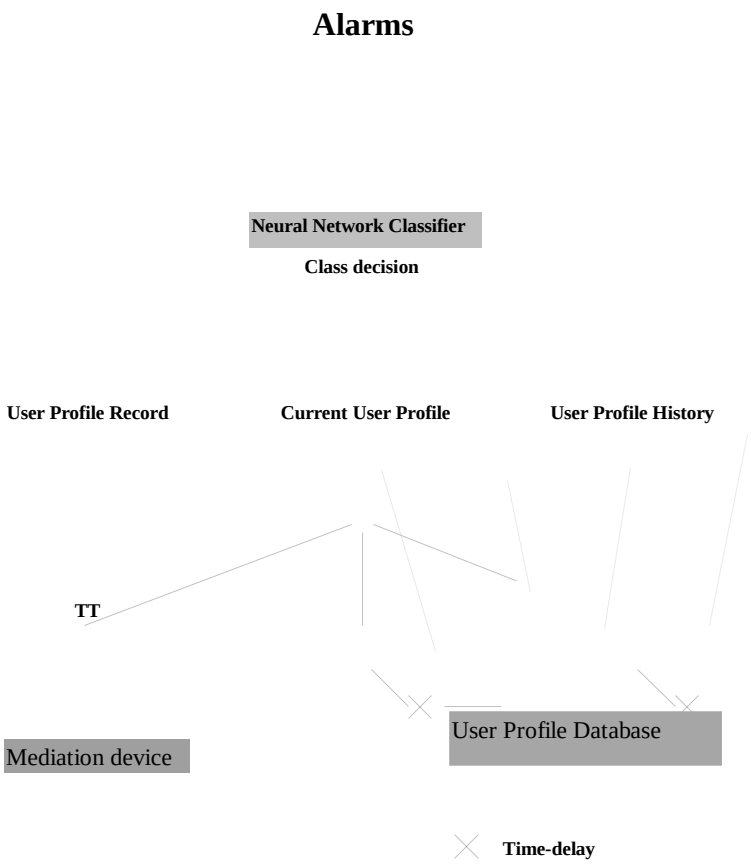


Figure 9-3 Run-time architecture of the neural network

At this point comes the issue of the management of the user profiles. The number of GSM subscribers, that is of user profiles, at Vodafone is currently in excess of 830,000. An upper bound of the speed at which the system would have to handle toll tickets would be about 30 toll tickets/ second. This would mean reading and writing 30 CUP and UPH per seconds to the hard disk. The fraud detection engine will not necessarily need to be able to sustain such speeds, but it definitely needs to be based on an architecture that could be scaled to attain such performances. The solution that has been chosen is to use a key-

Figure 9-2: Architecture for the training of the neural network

9.5.2 Run-time architecture

Let us now consider the run-time behaviour of the system. When toll tickets are produced by the mediation device (or this is an extra layer implemented in the hardfraud detection engine. The front-end extracts the UPR from the toll ticket. The CUP and UPH retrieved from the hard disk. The CUP is updated using the UPR and the UPH is updated first contains the list of IMSI, the second contains the list of CUPs and UPHs, when an IMSI is given to the database manager, it searches the key in the hash table. The UPR, CUP, UPH can then be input to the neural network classifier. The classifier will then output a class decision. A different strategy is used for storing the data. The database is flushed, for fast access.

If the classification is positive, i.e. a fraud is detected, that is the output corresponding to one class is greater than a threshold, an alarm is raised, the outputs and the user profiles are sent to a fraud investigation operator, and an audit trail is written to an alarm log.

9.6 Run-time simulation of toll tickets arrival

Although Vodafone is in the process of building a mediation device that will deliver the toll tickets in real-time, this system will not be available at the time of the demonstration of the prototypes. Therefore, the toll tickets available for testing will only be a batch file of toll tickets. In order to reproduce the real conditions of use of the fraud detection engines as closely as possible, it is necessary to develop an algorithm that will simulate the behaviour of the mediation device.

This can easily be done using an exponential distribution of the interval of time between the arrival of two successive toll tickets. The interval of time between two arrivals is a random variable Θ. The idea of an exponential distribution is that the process is memoryless. That is, the probability that no toll ticket will have arrived by time τ is equal to the probability that there will be no arrival up to time τ+ρ if we knew that no toll ticket had arrived by time ρ. The cumulative density function for such a process (that is, the probability that a toll ticket will have arrived by time τ) is as follows:

$$P(\Theta < t) = \begin{cases} 0 & t < 0 \\ 1 - e^{-lt} & 0 \le t < \infty \end{cases}$$

where l is a positive parameter defining the distribution. The mean arrival time is $E(\Theta) = {}^{1}{}_{l}$ and the variance of the arrival time is $s^{2}(\Theta) = {}^{1}{}_{l^{2}}$. A random variable with this distribution can easily be generated using a uniform random number generator and inverting the cumulative density function. If is a uniform random variable on the interval [0,1], then ${}^{1}{}_{l}\log(1-X)$ is an exponential variable with parameter l. To simulate the arrival of the toll tickets, we simply have to read a toll ticket in the batch file, compute a number r between 0 and 1 using a uniform random number generator, and then compute $= -{}^{1}{}_{l}\log(1-r)$. The toll ticket is then released to the fraud detection engine after a time interval Θ. The next toll ticket is then processed in the same way.

10. Neural Network Based Approach to Fraud Detection: Unsupervised Learning

We describe here a Neural Network system for detecting anomalous behaviour through the Unsupervised Learning of statistical User Profiles, generated by prototyping GSM Toll Tickets.

10.1 Introduction.

Neural networks have been shown to be particularly good at classification. If a problem can be transformed into such a classification task then there are many well established Neural Network techniques that can be applied to solving the problem. The most important stage in the development of such a system is how to represent input data and the way in which it is presented to the Neural Network.

When considering the task of detecting fraudulent activity in mobile telecommunications networks, the challenge is to find a suitable representation of Toll Ticket data to summarise mobile phone usage and form user behaviour profiles. The format of these profiles should suit the needs of the specific fraud engine. This section considers a profiling technique suited to Unsupervised Learning in Neural Networks. User profiles will be generated automatically by the classification of GSM Toll Tickets into one of a set of Toll Ticket prototypes. As Toll Tickets are classified, statistical information pertaining to the number of times a Toll Ticket prototype is excited, by the presentation of an incoming Toll Ticket, will be computed and stored in the form of a record, the length of which will be equal to the number of prototypes.

By considering two different time spans over the Toll Tickets, we generate two profile records for each user. The profile representing the shorter Toll Ticket span can be considered as representing the user's most recent activity. This we will call the Current User Profile (CUP). The longer span will create the User Profile History (UPH). Training can be seen as the presentation of clean profiles to the system to define the boundaries of acceptable behaviour, based on a differential analysis. The task of the system, after training, is to raise alarms when it is presented with profiles where the difference between the CUP and the UPH is outside the realms of normal usage. An alert status will be raised if the profiles are significantly different.

This system requires only clean (non fraudulent) data for training. This is advantageous because the fraudster has not yet caught up with GSM technology and few fraud scenarios are seen in practice. In addition, this system has the potential to detect new types of fraud as and when they occur. The system is able to do this because it is not being trained to recognise specific fraud scenarios, but rather altered or unusual usage.

10.2 The prototyping technique.

Prototyping is a method of forming an optimal discrete representation of a naturally continuous random variable. The processing of continuous random variables by discrete systems generally reduces empirical information. Neural Networks are capable of forming optimal discrete representations of continuous random variables through their ability to converge, by lateral interaction, to stable uniformly distributed states.

The maximum-entropy principal states that
the mapping of a continuous random variable X into a set of K discrete prototypes Q reduces the empirical information by the least amount if a uniform distribution $\{\mathbf{P}(q_i) \quad \frac{1}{K}, i = 1 \ldots K\}$*, corresponding to the absolute maximum* $(S_Q = \log K)$ *of information entropy, is assigned to Q.*

In [6], Grabec provides a way to extend this principal to multiple dimensions. When considering the set of all possible Toll Tickets, we clearly have a dimension to represent every parameter that we wish to include in the analysis. Each parameter in a Toll Ticket can assume a range of values and is thus itself a random variable. Grabec's technique will allow us to create a number of prototypes that dynamically and uniformly span the set of samples taken from the space of possible Toll Tickets. In this way we will be able to classify a new incoming Toll Ticket by the prototype that most closely resembles all its characteristics.

A user profile can then be constructed as a vector of counters representing the number of times each prototype has been excited by the presentation of an incoming Toll Ticket, for that user. In order to maintain the notion of a span over the toll tickets, we propose applying a decay factor to the vector of counters prior to incorporating information from any incoming Toll Ticket. By using two different decay factors, we will maintain profiles representing the two different spans over the Toll Tickets, namely the CUP and the UPH.
In order to generate a set of condensed Toll Ticket prototypes, we require a clean data set of Toll Tickets. These will be provided in large quantities by Vodafone and Panafon. The Toll Tickets are then condensed by extracting parameters that are considered relevant to the analysis. An iterative procedure is set up to dynamically distribute prototypes over this Toll Ticket sub-space through sampling the incoming stream of condensed Toll Tickets. This can most clearly be seen through defining the procedure for a one dimensional space, i.e. considering a condensed Toll Ticket consisting of one parameter. We first define the iterative procedure to calculate the change in the current value of the K prototypes Q;

$$\Delta q_l^{(i+1)} = B_l - \sum_{k \neq l} C_{lk} \Delta q_k^{(i)} \quad ; l = 1...K \tag{10.1}$$

starting with $\Delta q_l^{(0)} = B_l$.

We define matrix *C* as

$$C_{lk} = \left[1 - \frac{(q_l - q_k)^2}{2s^2}\right] \exp\left[\frac{-(q_l - q_k)^2}{4s^2}\right] \tag{10.2}$$

where $s \approx \frac{S}{K}$ approximates the standard deviation and S is the expected range of X. In practice, the results are not particularly sensitive to the choice of s and so a broad estimation of the range S of X will suffice.

Vector **B** is defined as

$$B_l = \frac{K}{N+1}\left\{(X_{N+1} - q_l)\exp\left[\frac{-(X_{N+1} - q_l)^2}{4s^2}\right] - \frac{1}{K}\sum_{k=1}^{K}(q_k - q_l)\exp\left[\frac{-(q_k - q_l)^2}{4s^2}\right]\right\} \tag{10.3}$$

Starting with $q_0(k) \quad X_k \; ; \quad k = 1,...K$. We calculate changes in Q on sampling a new condensed Toll Ticket from X. Using (10.1) we iterate calculating a vector of changes ΔQ based on the new Toll Ticket. Experimental results have shown that four iterations of (10.1) will be sufficient to stabilise the vector. The current vector of prototypes Q can then be updated by ΔQ.

The above procedure extends to the multidimensional case where condensed Toll Tickets consist of more than one parameter. The vector of prototypes becomes a vector of 'vector prototypes' - in other words a matrix. The process is repeated, by sampling condensed Toll Tickets until we find that the matrix $\Delta Q < e$, an arbitrary stability criterion. When this condition holds over a defined number of incoming condensed Toll Tickets, the matrix of prototypes can be considered to uniformly span the condensed Toll Ticket space.

One area which we plan to investigate is the adaptation of the prototyping technique to use a non Euclidean metric between prototypes as some condensed Toll Ticket parameters may be of more importance to the analysis than others. Also the relationships between toll ticket parameters could potentially be far from Euclidean.

10.3 Constructing a statistical user profile.

Once we have constructed a matrix of condensed Toll Ticket prototypes we can begin the task of building the user profiles. At this point, updating of the prototypes ceases. With a large enough data set, defining the space of possible condensed Toll Tickets, there should be little movement among the prototypes. Further more, any drifting of prototypes would occur over a long period of time. Due to the finite span of Toll Tickets forming the profiles, it is unlikely that this drift would be noticed.

As condensed Toll Tickets arrive for a user, each one is classified by one of the prototypes. A vector of counters keeps track of the number of times each prototype has been excited by an incoming condensed Toll Ticket, for that user. This vector of counters becomes the user's profile. A simple example is drawn below in Figure 10-1.

Subscriber	2	3	4	1	2	0	0	0	2	1

Figure 10-1 A user profile counting prototype excitations. 15 Toll Tickets have arrived and there are 10 prototypes.

10.4 Maintaining current and history profiles as probability distributions.

As mentioned in the previous section, we need to maintain the concept of two different spans over the condensed Toll Tickets to perform a differential analysis. We point out though that a differential analysis should also trap any frauds that would be detected through an absolute analysis due to their extreme nature. There is one special circumstance which we will need to cater for which is not having a history profile for a new user. It is envisaged that a new user will be assigned a default history profile based on the subscription tariff that was chosen. As soon as Toll Tickets start to appear for this new user, they will be incorporated into the user profile.

We propose to maintain the two profiles as probability distributions using two different decay factors a and b. When a Toll Ticket is presented to the system to update a user's CUP, each element of the CUP is multiplied by decay factor a. The entry in the profile corresponding to the prototype i, that was excited by the presentation of the incoming condensed Toll Ticket, is then incremented by an amount $(1-a)$. Updating the CUP in this manner will maintain the profile as a probability distribution. After updating the CUP, both profiles are presented to the fraud engine as discussed in the next section. Following presentation to the fraud engine, the UPH is updated using

$$H_i = bH_i + (1-b)C_i$$

where H_i and C_i represent the i_{th} element of the UPH and CUP respectively.

The exact value of the two decay factors a and b may prove to be critical to the success or failure of the system and will be determined by experiment. Figure 10-2 demonstrates the effect of decaying a CUP.

Level of profile after decay.										
Subscriber	1.94	2.91	3.88	0.97	1.94	0	0	0	1.94	0.97

Figure 10-2 Decaying a CUP by multiplying with decay factor.

By applying a multiplicative decay factor, any counter in the profile, corresponding to a prototype, once excited will never actually decay to zero. This is important because if a particular behavioural event occurs very infrequently, such as an overseas call for some users, this could be seen as a behavioural anomaly if the profile entry corresponding to it was zero.

10.5 The fraud engine.

The fraud engine will work in two modes, a training mode and a detection mode. In the training phase, clean user profiles, in the form of probability distributions, will be presented to the fraud engine. It is the task of the fraud engine to determine the difference between these distributions.

To perform this task, we will use a measure known as the Hellinger distance shown below in Equation (11.4).

where C and H are the CUP and UPH respectively. The Hellinger distance will always be a value between zero and two where zero is for equal distributions and two represents orthogonality. As clean profiles are passed to the fraud engine, in training mode, we determine the maximum value of d that was found $d_{threshold}$. This value represents the greatest difference found between a CUP and UPH and is essentially the most erratic behaviour seen in our clean training set of user profiles.

In detection mode the fraud engine again calculates d according to Equation (10.4). If the resultant value of d is greater than our threshold value computed in training then an alert status is raised proportional to $d - d_{threshold}$. It is anticipated that these alert statuses will be prioritised for investigation.

10.6 The system level design.

Potential systems architectures were discussed in 'Frameworks for Fraud Detection in Mobile Telecommunications Networks'[5]. This system is based on one of the potential system designs presented in that paper. The first implementation of the system will have a single component fraud engine. The task of the fraud engine will be to determine the likelihood that a fraud has occurred based on the CUP given the UPH. Unsupervised learning will take place off-line with clean sets of Toll Tickets. Due to the processing requirements of the real time system, a cache to virtual memory may be needed to swap the profiles of infrequent users out to disk retaining only the most recently active user profiles in memory. The first implementation of the system will not contain this feature and all profiles will be stored on disk. Figure 10-3 shows schematically how off-line training will take place.

Clean Toll Ticket arrives for user X

Backing Store.

Caching system.

Storage of User Profiles.

Extract features, forming Condensed Toll Ticket. Compute closest prototype to this Toll Ticket.

Present the CUP and the UPH to the Fraud Engine in training mode..

Apply CUP decay factor to CUP and update with prototype.

Apply UPH decay factor to UPH and update with prototype.

$$d = \sum_{l=0}^{K} \left(C_i - \sqrt{H_i} \right)^2 \quad (10.4)$$

Notice that the new prototype, generated by classifying the condensed incoming Toll Ticket, is not added to the UPH until after the fraud engine has processed the updated CUP. This is done so as to amplify the difference between the CUP and UPH.

In Figure 10-4, the system is shown in detection mode. The way in which the profiles are updated is the same as in training mode. The fraud engine now has the task of performing the vector difference operation (10.4) and computing any alert status reporting the information to the operator's terminal.

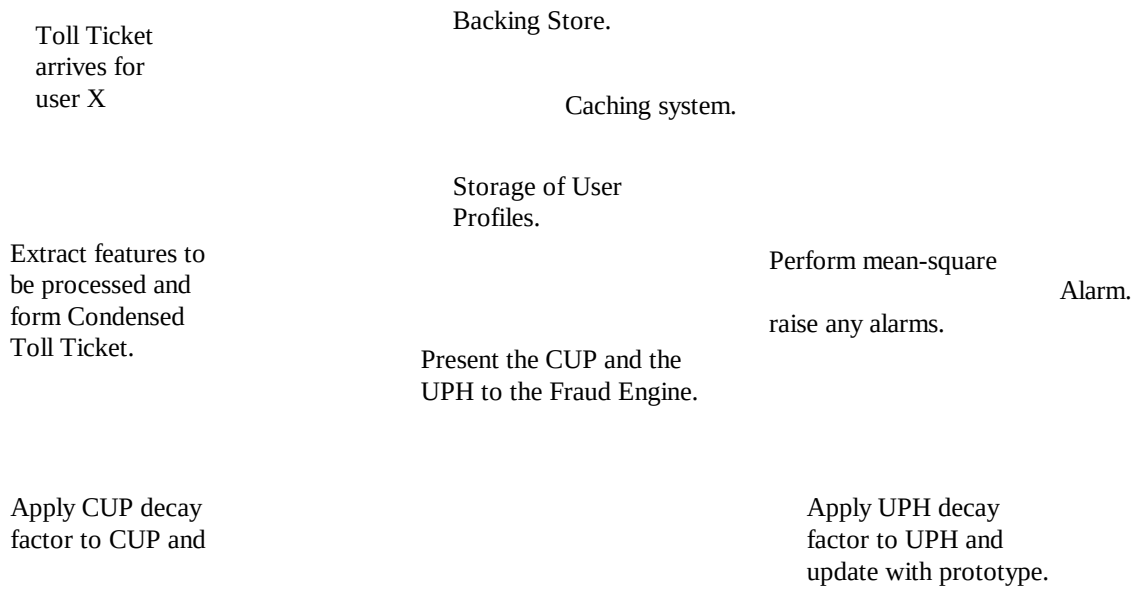


Figure 10-3 Off-line unsupervised training of the fraud engine.

10.8 Plans for later versions of the system.

In later versions of the system, we will

- Investigate other non Euclidean relationships between prototypes that may prove to be more meaningful to the analysis.
- Adjust the profiling technique so that not only one counter gets incremented in the user profile but all counters get incremented proportionally to this new non Euclidean metric.
- Introduce the cache system in order that a real-time implementation can be demonstrated.
- Improve the user interface to a common standard agreed within the work package.

10.9 Results.

An initial attempt at prototyping simulated one dimensional condensed Toll Tickets has been performed. In Figure 10-5, we see five prototypes of call start time being generated over 1000 incoming condensed Toll Tickets. In this situation only one iteration of Equation (10.1) is being performed. Call start time is being measured in seconds from midnight.

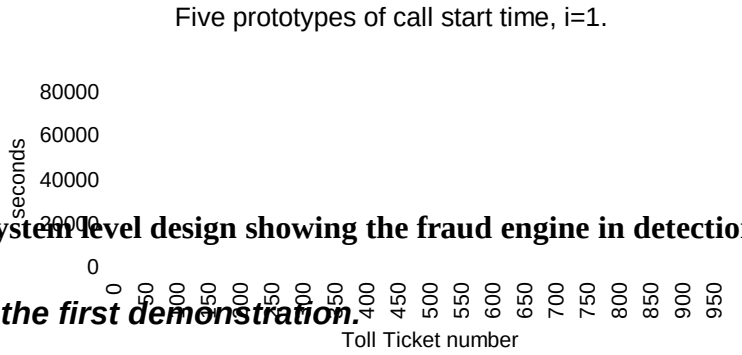


Figure 10-5

Notice that the spread of prototypes by Toll Ticket 1000 is a rough approximation to uniformity but not exact. In Figure 10.6, however, the number of iterations of Equation (10.1) has been extended to 10. This time the excitation probability distribution over the prototypes is more uniform.

Figure 10-4 System level design showing the fraud engine in detection mode.

10.7 Plans for the first demonstration.

Our initial implementation of the Neural Network system will be to

- Pre-train a fraud detection tool based on the above design using non fraudulent sets of GSM Toll Tickets provided by Vodafone and Panafon.
- Read in a new sequence of Toll Tickets, from disk, containing fraudulent examples and to output the results of the fraud engine in a simple ASCII format.
- Demonstrate the flexibility of prototyping and show the effect that adjusting the number of prototypes has on the overall performance of the system.
- Show what the effects of changing the decay factors on the CUP and UPH are. It is thought that the value of these parameters will be critical to the success or failure of the system.
- Only consider a Euclidean metric for distributing the prototypes over the space of possible condensed Toll Tickets.