An Approach to Rules based Fraud Management in Emerging Converged Networks

Jimmy McGibney & Seán Hearne
Telecommunications Software & Systems Group, Waterford Institute of Technology, Ireland
Tel: +353 51 302906; Fax +353 51 302679
Email: {jmcgibney, shearne}@tssg.org

Abstract – This paper examines the problem of managing fraud in emerging converged networks and presents work in progress on implementing a rules based fraud detection system for deployment in a testbed environment. This is done by examining the state of the art in telecoms fraud management and adapting this to emerging IP-based networks and services. Features of this fraud detection implementation include the use of flexible data formats and spreadsheet/workbook-based rules specification with the capability to apply arbitrarily complex rules.

Acknowledgement – This work is carried out as part of the *Converge* project, funded by the Irish Government under Strand III of the Technology Sector Research Programme.

1 Introduction

The ongoing convergence of voice and data traffic, together with migration of traditional telecommunications services to packet switched networks based on the Internet Protocol (IP), necessitates investigation of major issues related to service delivery, quality of service (QoS), usage based accounting, and security. Security issues include securing access to services, securing the accounting process and securing the communications themselves. Fraud management, the focus of this paper, combines aspects of accounting with aspects of security, and is in our view of great importance in an emerging environment where services are numerous, diverse and potentially short-lived and where users can expect quality guarantees but will need to pay for them.

Even in the circuit switched telephony environment, which arguably is well protected against fraud due to highly restricted access, revenue loss through fraudulent activity is often valued by service providers at between 3% and 8% of turnover [1], amounting worldwide to tens of billions of euros per annum. It is very likely that emerging IP-based services will provide even greater opportunities for fraud, due to the use of multi-layered, open protocols and their lack of built-in security mechanisms (with IPv4 at any rate).

Indeed, IP-based telecommunications networks will be susceptible to the same security problems that are inherent in any IP network. As access to IP networks is largely unrestricted and IP fraud may be performed from multiple points in the network simultaneously, the successful detection and management of this activity requires constant exchange of information between several network elements, devices, and interfaces, followed by much comparison and analysis.

This paper investigates techniques for gathering fraud indication data and processing it according to user-specified rules so that suspected fraud can be discovered and acted upon. A spreadsheet workbook-based approach is proposed for flexibly handling rules specification.

2 Problem statement

Fraud management is a very general issue, which we can define as including the following:

- *Fraud prevention*: the enforcement of strict access and usage controls to ensure that fraud cannot take place.

- *Fraud detection*: real-time or non-real-time observation of indicators (mainly service usage metrics) and determination of whether fraud is taking place or has taken place. This usually triggers some action, such as blocking access to the service or generation of a notification.
- *Fraud reduction*: recognising that fraud prevention is almost impossible in practice, ensuring that it happens rarely and that its effects are minimised. This usually requires real-time detection.

Fraud prevention (or at least minimisation) is an essential part of the design of services for deployment on emerging networks. Fraud prevention usually depends on the implementation of standard security facilities like authentication of identity and policy-based authorisation, and involves some kind of *a priori* processing. Significant work on securing access to charged services has been undertaken by the Authentication, Authorization and Accounting (AAA) Working Group of the Internet Engineering Task Force [2].

The central problem that we address in this paper is that of *a posteriori* processing in order to detect fraud – accepting that, despite the enforcement of strict controls, there remains the possibility of fraud. This processing takes place either while a service is being used or some time afterwards. Figure 1 shows the context of this fraud detection.

Inputs to this fraud detection process, discussed in more detail below, include usage data gathered from the currently active service, historical usage data, customer/user profile information and a set of fraud detection rules that are to be applied to this data. These fraud detection rules might be tailored to individual services, individual customers, or both.

Outputs from the fraud detection process can be one or more of:
- A Boolean value, indicating whether or not the current service usage is considered fraudulent;
- A quantitative measure indicating the likelihood of fraud;
- An updated customer/user profile, to be used as input to future fraud detection activity.
- Updates to historical data that is maintained on service usage

Handling of these outputs is beyond the scope of this paper. A determination of fraud would normally need to trigger some action, ideally in real time, possibly generating an alert within the service provider's existing operations and management system.

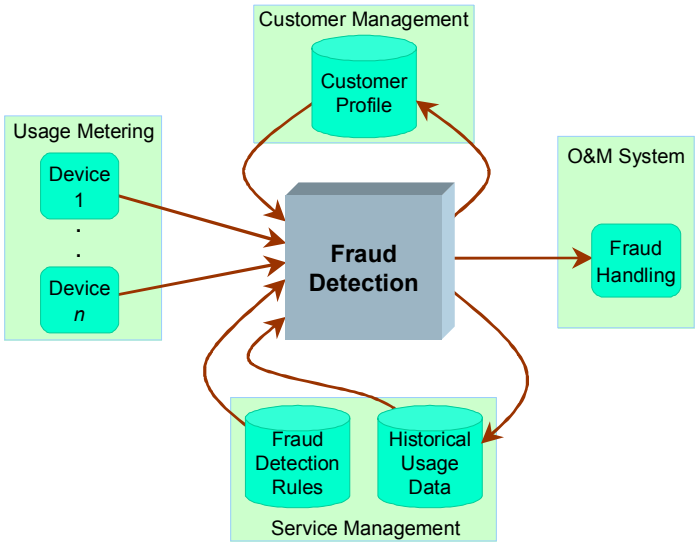

Figure 1: *The Fraud Detection process in context*

3 Design Approaches

There is a significant body of published work on telecommunications fraud detection, mostly addressing circuit-switched second-generation mobile networks. This section seeks to adapt this work, and some additional relevant work on network intrusion detection, to emerging converged networks.

The main design issues that we can identify for a fraud detection system are the following:

- Identification and management of fraud indicator data
- Collection and representation of usage data
- Processing of activity data – i.e. the rules engine
- Specification and maintenance of rules
- Deployment (distribution of logic, performance, scalability)

Each of these is now discussed individually.

3.1 Identification and management of fraud indicator data

The first step in designing a fraud detection system is to analyse the type of data that is available for monitoring so that fraud indicators can be identified. There are two major categories of such data. Firstly, data is normally available for each individual service usage, often summarised in the form of toll tickets or call detail records (CDRs) that are created for billing purposes. This data may or may not already be *rated* (i.e. have charges applied). Service usage data may also be available from various kinds of log files that are maintained for performance monitoring, or specifically for the purposes of fraud detection. Secondly, a fraud detection system should have access to persistent data that is not directly related to the "current" service usage. This includes data from customer management systems as well as historical usage data, which can be potentially vast. A major challenge is to store and access this kind of data as efficiently as possible.

As mentioned, indicators of fraud are usually available from service usage or network monitoring records. Various kinds of usage-based indicators are discussed by Burge et al in [3], including *usage indicators*, related to how a service is used (e.g. frequency of use), *mobility indicators*, related to the mobility of the originating device or person, and *deductive indicators* that arise as a by-product of fraud (e.g. overlapping calls). Indicator data may be available immediately after a service usage has begun, ideally enabling fraud detection while the service is being used, or only after service usage has been completed, as is the case with circuit switched telephony's CDRs that are often intended for batch processing.

Fraud indicators are also available from customer management systems – i.e. individual user profiles. Data from a particular service usage instance can then be combined with the user's personal profile to indicate the likelihood of fraud. It is envisaged that profiling of individual users can greatly increase the power of fraud detection. This may be done by building up a picture over time of each customer's service usage, and perhaps other information, like the customer's credit rating. Cahill et al in [4] argue for the tracking of each account's own activity, and propose the use of a *signature* to represent this in an efficient way. This signature is kept current by the use of a *decay mechanism*, whereby recent data is afforded greatest significance. The extent to which separate data is maintained on each individual user has an effect on performance and scalability. At one extreme, we could have simple classification of users, while at the other we could have highly individualised data, perhaps by allowing customers to define their own profiles. This has the advantage of actively involving the user in fraud minimisation.

3.2 Collection and representation of usage data

As mentioned, the records most commonly monitored for fraud indicators in circuit switched telephony are CDRs. A CDR usually contains information about a completed telephone call or call attempt and is used for billing purposes.

In our fraud detection design, we choose to represent service usage data with Internet Protocol Detail Records (IPDRs), in accordance with the IPDR organisation's Network Data Management – Usage (NDM-U) specification [5]. The NDM-U specification is a generalisation of the CDR idea, and it can be used for all kinds of diverse services. The IPDR organisation is an industry consortium, founded by several prominent vendors of management solutions for IP-based networks and services. The main objective of the IPDR organisation is "to define the essential attributes of information exchange between network elements and services, operation support systems and business support systems" [6]. The specifications are based on the core functional roles and interfaces of the TeleManagement Forum's Telecom Operations Map (TOM) [7]

There are several reasons for adopting the NDM-U specification for fraud management for IP-based networks and services:

- IPDR NDM-U is an industry-wide specification. CDRs are generated by telephone switches, usually in a format that is specific to the switch vendor. Though there has been some attempt to standardise formats (e.g. GSM's Transferred Account Procedure for exchange of usage data between roaming partners), billing and fraud detection systems have generally had to be tailored to specific formats.
- The IPDR structure specifies a generic, flexible record format for exchanging usage information in a multi-service environment. CDR formats are generally quite static.
- IPDR provides an extension mechanism so that additional, optional, usage metrics may be exchanged for a particular service, or even a particular service usage instance.
- IPDRs can be used for exchanging any kind of usage data, not just data on completed calls. For example, IPDRs can be generated periodically while a service is being used to enable near real-time billing (and fraud detection).
- IPDRs are self-descriptive and human-readable, based on eXtensible Markup Language (XML), allowing for more straightforward integration into diverse systems. Figure 2 shows a sample IPDR for a telephone call.

A potential drawback with the IPDR model is that, in common with most self-describing text-based specifications, it is not the most efficient way to represent data. Efficiency can be improved however with the use of native XML databases as well as XML compression.

```
<?xml version="1.0" encoding="UTF-8"?>
<IPDRDoc xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
docId="SMH_20030224143129+0100_24194"
xmlns="http://www.ipdr.org/namespaces/ipdr" IPDRRecorderInfo ="recorder.acme.net">
    <IPDR xmlns="http://www.ipdr.org/namespaces/ipdr">
        <seqNum>3462</seqNum>
        <IPDRCreationTime>2003-04-28T08:53:05Z</IPDRCreationTime>
        <serviceType>fixedLineTelephony</serviceType>
        <serviceProviderID>AcmeTelecom</serviceProviderID>
        <ANumber>050679390</ANumber> <!- Calling Party -->
        <BNumber>1891110110</BNumber> <!- Called Party -->
        <startTime>2003-04-28T08:50:00Z</startTime>
        <duration>185</duration> <!- Duration in Seconds -->
        <CE xmlns="http://www.ipdr.org/namespaces/ipdr">
            <charge>0.062</charge> <!- €.02 per minute -->
        </CE>
    </IPDR>
</IPDRDoc>
```

Figure 2: *Sample fixed-line telephony IPDR*

Of course, not all data of interest for fraud detection will be available in IPDR format. Thus, one of the required subsystems of our fraud detection system is a *mediation* component that is responsible for pre-processing data with appropriate correlation and aggregation, and presenting IPDRs to the rules engine.

3.3 Processing of activity data

Published work on fraud and intrusion detection generally takes one (or both) of two approaches: rules based detection and/or "intelligent" methods using neural networks, data mining or case based reasoning. These approaches differ in their ease of use, adaptivity, complexity and performance. In this paper, we focus on rules based techniques.

With rules based detection, usage data is verified against specified rules. These rules may be *absolute* or *differential*. The former are based on simple thresholds, which may or may not be customer-dependent. The latter are based on observed statistical anomalies, the identification of which can be based on customer profile, time of day or other factors. A statistical anomaly occurs when there is a perceived difference between observed behaviour and "normal" behaviour. Many computer intrusion detection systems are based on anomaly detection.

A good fraud detection system needs to have a high detection rate while minimising false positives. If too few frauds are detected, then the system is not very effective; if there are too many false alarms, then they cause time wastage and tendency for people not to take them seriously. A notable problem is what is known as the *base rate fallacy*, a consequence of the probabilities involved, which shows that it is very difficult in this type of system to achieve a high rate of detections with a low rate of false alarms [8].

Rules based fraud detection is usually implemented by some kind of predicate logic that works on input data. The main implementation issues for rules based detection are mediation of this input data to some standard format, choice of rules engine, and provision of flexible tools for specification of arbitrarily complex rules.

3.4 Specification and maintenance of rules

Another major design question for rules-based fraud detection systems is how rules are stored and edited.

If fraud detection rules are static then their effectiveness is reduced, firstly as this implies that they cannot be tailored to one-off or rapidly changing services, and secondly as perpetrators of fraud tend to get to know the rules and develop workarounds. Thus it is paramount that rules are easily editable, and are highly customisable, either per-service or per-user. An ideal scenario is where the customer is actively involved in rules specification (e.g. "I rarely make international calls, and when I do they're almost always to France").

Furthermore, for maximum effectiveness, fraud detection rules should be able to be dependent on *any* input data – i.e. arbitrary choice of IPDR fields and other input data, and it should be possible for these rules to be almost arbitrarily complex. Formally, we can write this as:

$$\text{DetectionResult} = f(\text{IPDR fields}, \text{HistoricalUsageData}, \text{CustomerData}),$$

where f is an arbitrary, non-linear function.

The challenge for the implementation of a fraud detection system is to define a representation of rules that is flexible and user-friendly, but sufficiently powerful to permit the most complex of rules to be captured. Specifically, the representation format must:

- Be easy to manage (by non-programmers);
- Permit arbitrarily complex formulae;
- Allow various input formats.

Motivated by experience with rating engine development [9], we propose to satisfy the above requirements by providing the user with a spreadsheet-like user interface. Spreadsheet logic, such as that of *Microsoft Excel*, is very flexible and tends to be familiar to personnel in finance and revenue assurance functions. As part of service design, usage data is identified to be placed into certain cells in a worksheet, and other cells are for the user to enter formulae or conditions. Sample usage data can be entered for testing purposes, but actual usage data is taken from IPDRs, etc., and placed into the specified positions at runtime, and the rules are executed automatically.

3.5 Deployment

Deployment issues include distribution of processing, storage (of rules, usage data & customer profiles), performance optimisation, and cost. A crucial consideration is whether to centralise rules processing, or to have this processing take place close to the edges of the network. The latter option may offer superior performance though, if the software architecture allows for it, centralised processing can match this by utilising load balancing techniques. The main advantage of detecting fraud close to individual devices is that it may be easier to trigger a reaction (e.g. cut off the device). The main advantage of centralised processing is increased scope for correlation of data from different locations.

The method used for user profiling has a significant impact on performance, storage and memory requirements. At one extreme, we can have simple *classification* rather than individual profiling of users. This minimises the quantity of data to be stored for each customer (there may be very many) and may also improve performance by, for example, allowing per-customer rules to be stored in memory rather than on disk. At the other extreme, we may have highly individualised profiles, for example by allowing each user to define his/her own. This approach has the advantage of actively involving the user in fraud minimisation. The system could even be configured so that the user receives a personal notification when certain events occur that might point towards fraud.

4 Implementation

Work is in progress on building a rules based fraud management system. This work builds on an existing accounting implementation [10] that applies charges to IPDRs making use of widely understood spreadsheet logic for rule specification and processing. *Microsoft Excel* is currently being used for the rules engine, though the architecture is sufficiently flexible to allow use of alternative spreadsheet engines. It is possible for example to have users specify rules with *Excel*, but to have rule processing carried out by a potentially more efficient engine, such as the *Formula One* ActiveX control [11].

Another feature of the implementation is generation of alerts when fraud is suspected, and notification of these alerts to relevant network equipment.

5 Conclusions and Future Work

Our work so far has highlighted the problem of fraud detection in emerging converged networks, has examined various solution approaches and has proposed an initial design for a fraud detection system. Following our software implementation, significant experimental work is required in a testbed environment to evaluate approaches to rules specification and a variety of issues like cost, performance and storage requirements. A variety of different deployment scenarios need to be considered in this.

As well as examining the performance of our rules specification and management system, there is significant need for more work on fraud detection techniques themselves, especially their adaptation to emerging multi-service networks. As well as rules based detection, experimental work is required to validate alternative techniques, like the use of neural networks and other intelligent systems.

More generally, in light of the convergence of telecommunications networks and general computer networks, it is reasonable to expect the disciplines of telecoms fraud detection and network intrusion detection to merge into a single field of study – perhaps called "anomaly detection".

References

[1] International Engineering Consortium, Tutorial on "Fraud analysis in IP and next-generation networks", http://www.iec.org/online/tutorials/fraud_analysis

[2] Internet Engineering Task Force, Working Group on Authentication, Authorization and Accounting, http://www.ietf.org/html.charters/aaa-charter.html

[3] P. Burge, J. Shawe-Taylor, C. Cooke, Y. Moreau, B. Preneel, C. Stoermann, "Fraud detection and management in mobile telecommunications networks", *2nd European Conference on Security and Detection*, London, April 1997

[4] M.H. Cahill, D. Lambert, J.C. Pinheiro, D.X. Sun, "Detecting fraud in the real world", *Handbook of Massive Datasets*, Kluwer Academic Publishers, pp. 911-929, 2000

[5] IPDR Organisation, 2002, *Network Data Management – Usage For IP-Based Services*, *Version 3.1.1,* October 2002. Available at http://www.ipdr.org

[6] C. Ryan, W. Donnelly, E. de Leastar, J. Cloney, "Value-based billing in a 3G IP services environment", 6th World Multiconference on Systemics, Cybernetics and Informatics, Orlando, July 2002.

[7] TeleManagement Forum, *Telecom Operations Map*, GB910, Version 2.1, March 2000.

[8] S. Axelsson, "The base-rate fallacy and the difficulty of intrusion detection", *ACM Transactions on Information and System Security*, August 2000.

[9] E. de Leastar, J. McGibney, "Flexible multi-service telecommunications accounting system", *International Network Conference 2000*, Plymouth, July 2000.

[10] J. Brazil, E. de Leastar, C. Ryan, M. Ó Foghlú, "Workbook approach to algorithm design and service accounting in a component-orientated environment", *IEEE Workshop on IP Operations and Management*, Dallas, October 2002

[11] Tidestone Technologies Inc., Formula One, http://www.tidestone.com/