Combining fraud and intrusion detection - meeting new requirements -

Håkan Kvarnström[1,2] **Emilie Lundin[1]** **Erland Jonsson[1]**

[1] Department of Computer Engineering
Chalmers University of Technology
SE-412 96 Göteborg
{emilie, erland.jonsson}@ce.chalmers.se

[2] Telia Research
SE-123 86 Farsta
hakan.k.kvarnstrom@telia.se

Abstract

This paper studies the area of fraud detection in the light of existing intrusion detection research. Fraud detection and intrusion detection have traditionally been two almost completely separate research areas. Fraud detection has long been used by such businesses as telecom companies, banks and insurance companies. Intrusion detection has recently become a popular means to protect computer systems and computer based services. Many of the services offered by businesses using fraud detection are now computer based, thus opening new ways of committing fraud not covered by traditional fraud detection systems. Merging fraud detection with intrusion detection may be a solution for protecting new computer based services. An IP based telecom service is used as an example to illustrate these new problems and the use of a suggested fraud model.

Keywords: intrusion detection, fraud detection, modelling, abuse

1. Introduction

Fraud in telecom and datacom services causes a substantial annual loss of revenue for many companies. Fraud can in this context be defined as a deliberate act of obtaining access to services and resources by false pretenses with no intention of paying. For telecom services, such as telephony, signalling abuse (e.g. blue-boxing) has been successfully used to place free telephone calls worldwide. Although improved technology such as digital switches has made signalling abuse difficult to exploit, new technological trends such as the introduction of mobile telephony induce new categories of fraud. Subscription fraud, call selling and cloning handsets are examples of the types of fraud that exist in mobile services. Abuse often utilizes a valid customer's service subscription to conceal the crime. A more elaborate discussion of fraud and telecom crime can be found in [3][4][5].

Telecom companies, insurance companies, banks and other businesses, have been studying fraud and fraud detection for many years and have probably spent more time and money on this than the research community. However, most of their efforts do not reach beyond the limits of the companies and have not been available to the public research community. Still, a number of published papers on the subject are available, most of which focus on detection methods. AI methods are the most common in these studies, including different kinds of neural networks, data mining, case based reasoning etc. There are also statistical methods for user profiling and combinations of methods [7]. Often these papers include experiments on fraud data collected in an authentic system, such as mobile phone call records [11], data from an insurance company [7] and financial transactions [1]. Most papers do not measure the actual gain from using their fraud detection method but in [2] a cost model have been used to evaluate the classification of tax declarations. Some papers, e.g. [3], [4], [5] and [10], describe the current fraud situation that telecom companies

face, and well-known frauds. However, these papers do not discuss any details of the detection process or any organised fraud model.

A great deal of research has been done in the area of intrusion detection (ID). The detection methods studied, e.g. [6], are in many cases almost the same as those used in fraud detection. More work seems to have been done on intrusion models, e.g. [9] and [15], than on fraud models. We found very few cross references between research papers in the two areas and none of them discuss effects of combining the areas.

This paper addresses the new problems that are appearing in IP based services (Section 2) and proposes a new fraud model (Section 3) that can be used in the detection process. In Section 4, our fraud model is applied to an IP based gambling service. We also discuss detection techniques and requirements for combining fraud and intrusion detection (Section 5). Some conclusions are given in Section 6.

2. Fraud detection for IP based services

2.1 Towards IP based telecom services

The increasing use of IP, i.e. the Internet Protocol, in telecommunication raises new problems. New telecom services are being rapidly developed today and many of these are heavily based on datacom services (e.g. IP). The market penetration of Internet and the increase in e-commerce applications leads us to believe that new types of fraud are to be expected in the near future. Many of these services, such as online gambling and media-on-demand services, require payment transactions susceptible to fraud. Further, the business model for IP based services differs to a large extent from traditional telecom services. Several different types of actors interact with each other to offer these services. In the traditional case, the main point of interest when looking for fraud is the interaction between the customer and the telecom operator. The new business models involve business relations between several actors (e.g. network, service and content providers), all of which must be taken into consideration. Any interaction where one actor provides a service and another actor pays for it, may be a source of fraud.

2.2 Fraud detection - traditional vs IP based

Traditional fraud detection (FD) applications for telecom services are often tailored to telephony applications. In future services, new technical components will be used for which traditional fraud detection techniques will not be directly applicable. It is evident that new methods of fraud detection must be developed for future telecom services. As new services are more computer and IP network based, there is reason to consider not only traditional fraud detection but also more general computer intrusion detection methods. Fraud detection is often very application specific and aims at detecting ongoing frauds or situations that precede frauds, i.e. events that lead to a person or company being able to profit at the expense of the company providing the service. Intrusion detection aims at detecting computer misuse, i.e. any unauthorized use of a computer system. This includes attacks against service availability and sabotage that are equally as important to detect. Computer intrusions are likely to be a part of the fraud model for new services.

There are some general differences between traditional telecom services and new IP based services. One positive thing about the differences is that some traditional frauds will die out. This can occur because the vulnerabilities exploited in the fraud do not exist in the new services or because the billing model has changed in a way that makes the fraud either impossible or unprofitable. Unfortunately, there is reason to believe that a larger number of new frauds will crop up than will disappear. Many of the old frauds are still applicable and as the Internet gains users' trust and acceptance as a market place, the value and volume of payment transactions will increase. Below are some general problems we expect to find in the new services:

- Values involved in the services are probably greater, which makes the motive for fraud even more obvious. This also means that the actors involved can lose more money.
- More actors are involved, which makes the services more complex. It is always harder to perform a fraud investigation when several actors must cooperate and data must be collected from several different companies. There are also more candidates for performing frauds.
- The services are more dynamic and probably have a shorter lifetime. This means that fraud management systems must be more flexible and have the capability to be efficiently implemented in a short period of time.
- New technical components will be used, and new weaknesses will thus be found. Many new components will probably be more insecure than old ones. For example, the IP protocol has many weaknesses that are not found in the network protocol used for telecom services today.
- No standardized billing information exists in the new services. The main source of information for fraud detection has traditionally been the call detail

record (CDR), which is not applicable to the new services in its current form. Either a new type of standardized billing record must be developed or fraud detection must be adapted to different data for each application.

In conclusion, it can be stated that the new IP based services are more complex and that the fraud management for those services is likely to be more problematic.

2.3 Actors

The expected actors in IP based services are described below. Other actors may appear for other types of services and some will probably disappear, but the list is still applicable for a large group of different services.

Customer. The customer is the end user of the service, i.e. the receiver of service content. In the traditional fraud management system, the customer is the main subject of interest in the search for fraud.

Content providers. The content provider is the actor that delivers the content that finally is delivered to the customers. For example, this may include video films, music tracks and other media suitable for digital distribution such as books, magazines and advertisements.

Service providers . The service provider bears the main responsibility for the service. Service providers come in many different shapes and are difficult to define. For example, a company providing a WAP/ SMS service or a web hostel would qualify as a service provider. In some cases the term sub-service providers is used to denote a service provider that offers a specific service. Examples of sub-service providers are:

Application service providers (ASP) . A service provider offering (remote) access to certain applications such as word processing, business systems (e.g. ERP) or other applications normally installed on the user's local computer. One of the key advantages of using an ASP is that software upgrades and maintenance are handled by the service provider, which gives large scale advantages and thus brings down the cost per user.

Internet Service Providers (ISP) . Internet Service Providers offer access to the Internet by some form of access network. A dial- up service is one commonly used form of access.

Network Providers . The network provider operates the underlying network (backbone) that connects different actors. Owing to the heavy investment cost of establishing a backbone, typical network providers are large telecom companies.

Access network providers . The access network is the network to which the customer is physically connected. For simple modem dial-up services, the PSTN network is used for access. For broadband services, cable-TV(CATV), xDSL or fiber connections may be used for access. The access network is usually connected to the network providers and one or more service providers.E-payment provider (EPP)

E-payment providers serve as mediators in business transactions. An e-payment provider often acts as a trusted third party known to both parties in the business transaction. The buyer knows that he will receive his products once the payment is complete at the same time that the seller knows that he will receive payment as soon as the product is delivered to the customer. Often, an e-payment provider handles transactions of relatively small value (micro payments).

Non-customers . The non-customer is a person or company not registered as a user of the service. He is an important actor in the fraud model. The non-customer must be taken into account in all business relations between the actors.

Each business relation in which money flows from one actor to another is a potential point of fraud. Figure 1 shows an example of how money and services may flow between the actors. All of these relations do not exist in all types of services. A telecom operator may take on several roles, e.g. as both network operator and service provider. This means that for specific services in which this is the case, it is not necessary to consider all relations between possible actors.

2.4 Fraud indicators

Fraud indicators are facts about service usage that may indicate fraudulent behavior. Some indicators that can be used for telephony fraud detection are long duration calls, high call volume and hot destinations. Fraud indicators are often not sufficient by themselves but should be used in combination with other indicators, time windows and threshold variation by customer group and/or time of day.

These indicators are often found after the service has been used for a long time and it takes a great deal of expertise or statistical material to tune a "thresholds" for the indicators. Most of the known fraud indicators for telephony services are not directly applicable in the new IP based services. A reason for this is that they are based in most cases on information found in the CDR, which does not exist in the new services. Many of these indicators are also very application specific.

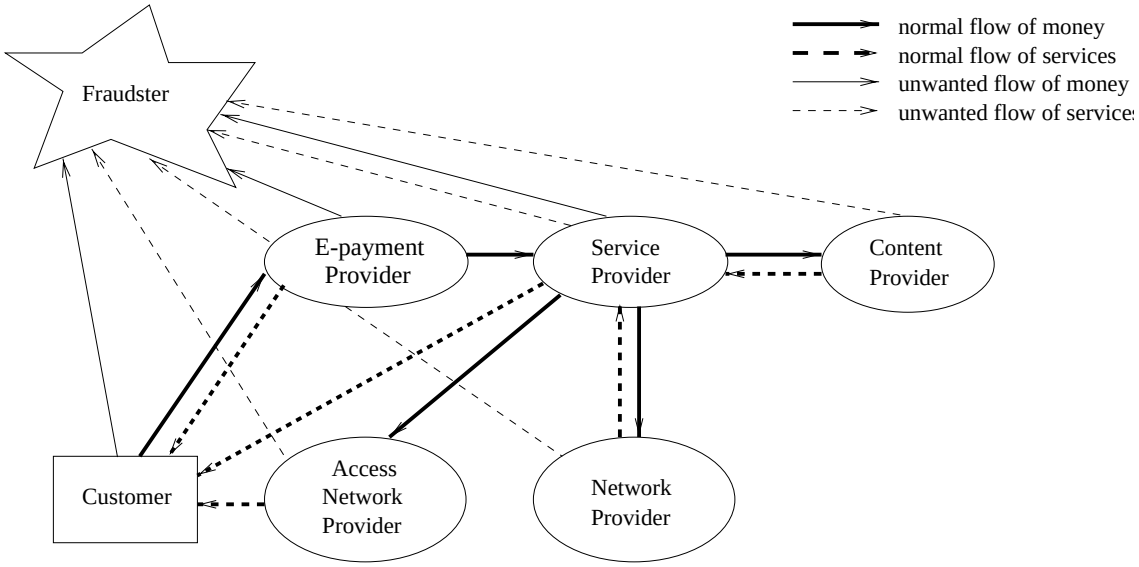

FIGURE 1. Relations between actors

Therefore, there is a need for development of new fraud indicators. Preferably, these indicators should be more general and applicable for a larger group of services. A requirement to make this possible is to standardize the format and contents of collected data.

3. A fraud model

It is obvious that it is necessary to have a clear definition of what constitutes a fraud in order to be able to successfully design a detection system. As mentioned above, while much work has been done to categorize and model intrusions, this is not the case for frauds. Some papers, e.g. [3] and [10], present frauds but not in a structured way. The purpose of these papers seems mainly to be to present existing, application specific frauds for telecommunications networks and mobile telephone networks. Their usability in designing fraud detection systems for new types of services is limited. This section suggests a model and way to categorize frauds.

3.1 The model

The objective of our fraud model is that it should be helpful in the process of designing fraud detection systems. The main idea of the model is to define groups of fraud characteristics important for fraud detection. The fraud can then be described by its characteristics, which determines what kind of "detection rules" can be applied. The higher levels of the model deal with general characteristics and the lower levels deal mostly with more application specific characteristics found only in a few services.

A fraud detection system can then be built in modules, where "detection rules" for frauds with certain general characteristics can be included in a standard module and modules with application specific "rules" can be added as needed. A high level "rule" can, for example, check whether the amount of money coming in to the service provider corresponds to the amount of services provided. A low level application specific module can contain rules applying to services using a specific billing model or technical component, e.g. rules for pay-per-view services or IP multicast services.

The fraud model is hierarchical, with more general parameters in the higher levels and more application specific parameters in the lower levels. The levels are presented in Figure 2. The highest level in the model is the general fraud goal, which applies to almost every commercial service. The second level deals with the actor involved in the fraud. The third level deals with the service specific risks that makes the service a target for fraudsters. Parameters on this level concern the content of the service, what billing models are used and vulnerable components. The last level deals with the technique used to accomplish the fraud. Here we must look at the "implementation" of the components used in the service, both physical and logical. Fraud indicators can be defined on the basis of a selection of fraud characteristics. Some indicators may include knowledge of characteristics on just one level of the model, but knowledge on several levels is often necessary. For example, the fraud technique used to commit a specific fraud may be to attack the network protocols. Then, a fraud indicator based on knowledge of only the fraud technique may be to look for deviations from normal behavior in the network traffic. A fraud indicator using knowledge from several levels may be

that the number of bytes of media delivered to the customer does not correspond to the service provider's billing log. Fraud indicators using characteristics in the higher levels of the model may be added without great knowledge of the target service, but indicators using characteristics in the more application specific layers require detailed knowledge of the service.

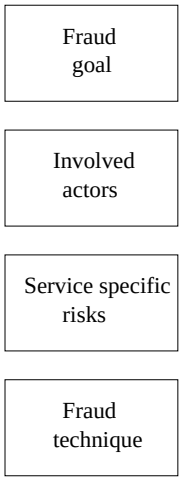

FIGURE 2. Fraud model

Fraud detection may be used by any of the actors involved in a service. The service provider is presumably the actor most interested in using fraud detection because he is involved in several sensitive business relations.

3.2 Fraud goal

The following is the dictionary definition of fraud: "Fraud is the intentional deception or misrepresentation that an individual knows to be false or does not believe to be true and makes, knowing that the deception could result in some unauthorized benefit to himself/herself or some other person". In telecom services the idea of a fraud is often that a customer wants to use a service without paying. However, there are also other types of fraud we want to include in our fraud model. Especially in the case of IP based services there are events we wish to detect that are not classical frauds but merely events of misuse that would be expected to happen in a computerized system. This includes sabotage and unauthorized use of resources. We assume that the following four classes of general fraud goals will cover what we wish to detect in a future fraud detection system.

1. Evasion of payment for service.
2. Service Level Agreement (SLA) violations. This means that an actor does not deliver contents of the agreed quality.
3. Improper use of service content, resources or procedures. This includes e.g. use of services in a unintended way and unauthorized use of resources, such as networks.
4. Sabotage of service.

A fraud detection system designer should not forget any of these classes when he considers what frauds can be expected for a specific service. The number of fraud indicators that can be defined by using only knowledge of the characteristics found on this level is limited, since not even the actors involved are known here. Something more than the fraud goal must probably be known about the fraud to make it possible to detect.

3.3 Involved actors

The second highest level in the model defines the actors involved in the fraud. The actor that commits the fraud is especially important, but it is also interesting to define the actors that may be affected by the fraud.

By studying the business relations in the service and the general fraud types, we can make reasonable predictions about which actors are likely to commit fraud and we can also analyze which actors will be affected by different frauds. There are six different actors in the gambling service described in Section 4.1. All these actors may have a motive for committing fraud. From the knowledge of fraud type and actors involved, it may be possible to design very general, and therefore reusable, fraud detection rules.

3.4 Service specific risks

The third level is the service specific risks. What is examined on this level is what it is that makes it interesting to commit fraud in this service. As mentioned before, this depends on the specification of the service, e.g. what billing models are used, the content of the service and the type of components used.

A specification of the service is required to define fraud characteristics on this level, but not necessarily an implementation. During the design phase of a service, we can predict frauds by studying the specification of the service and make comparisons with other services with similar contents, billing models or components. If the study shows that high risk frauds that are difficult to detect may appear, a revision of the service specification may be appropriate.

3.5 Fraud technique

The lowest level in the model is the technique used for committing fraud. The technique depends on what

components are used in the service. Fraud techniques can be divided into two sub-categories technical frauds and social engineering frauds. A technical fraud involves attacking a technical component, while a social engineering fraud attacks service procedures controlled by employees. The service procedures can be viewed as logical components. Of course, a combined fraud attacking both types of components is possible.

Some technical components expected to be used in most services are:

- Authentication mechanism
- Customer database
- Application server
- Network and network protocols
- Billing data generation and storage
- Service content storage

Examples of procedures controlled by employees:

- Customer registration procedure
- Customer support

To predict actual fraud techniques, we must have detailed knowledge of the service components and how they are implemented. Fraud indicators on this level are connected to specific service components and can be reused only for other services using the same components. If frauds are to be detected on this level, intrusion detection may be useful because many of the technical components used in an IP based service are general computer system components.

4. Application of the fraud model

This section describes an IP based service and some possible frauds that may occur in this service. The frauds are then studied using our fraud model.

4.1 Gambling service

To give an example of a possible new IP based service, we here describe an online gambling service.

4.1.1 Service description

Online gambling such as Lotto is sold through a network of retail stores that sells gambling services. Figure 3 shows a fictitious gambling service that is provided to customers. Several types of actors are involved, such as a content, network, Internet, and service providers.

Content providers may offer a variety of games such as sports betting (e.g. horse races), lottery (e.g. lotto) and gambling (e.g. cards, roulette). In this fictitious example the content provider (CP) specializes in various card games that are offered to the public. The CP allows customers to place bets by utilizing a network of service providers such as online retails stores, Internet gambling sites and affiliates that resell the gambling service (possibly branded under a different name). In addition, customer credit checks and payment transactions are handled by a third party specializing in online payment. Depending on the type of service provider, the payment scheme may vary, ranging from micro payments to credit cards and prepaid accounts. The advantages of using an e-payment provider (EPP) is that customers do not have to release their credit card information or establish accounts at a large number of service providers. Once a customer is registered to an e-payment provider, he can access all services offered by service providers using that EPP. At the same time, the service providers do not have to establish processes for customer billing and verification of customers credit status.

Customer transactions. A typical service transaction involves a series of steps: (1) The customer contacts a service provider that offers a card game; (2) The user identifies himself and is authenticated using some authentication mechanism (e.g. smart card, one-time password etc.); (3) The service provider verifies that the user (still) has a good credit standing by contacting the payment provider; (4) The customer is then allowed to enter a card game where bets are placed. All losses and wins are debited from/credited to the user's account.

Service provisioning transactions. The transactions between the content provider, service provider and payment provider are handled by a clearinghouse function. 50% of the revenue/loss goes to the content provider, 40% goes to the service provider and 10% goes to the payment provider. The risk is equally divided between the parties. The clearinghouse function is managed by the service provider. At regular intervals (daily), payment transactions are made to clear withstanding balances between the different actors.

4.1.2 Fraud scenarios

The possible frauds in this model are many. Here we present a few known frauds that are applicable and may occur in the gambling service. Starting with the *customer* related fraud cases, the possible frauds include:

Subscription fraud. A user registers himself under false identity and manages to access the gambling service. If the customer loses money, he has no intention

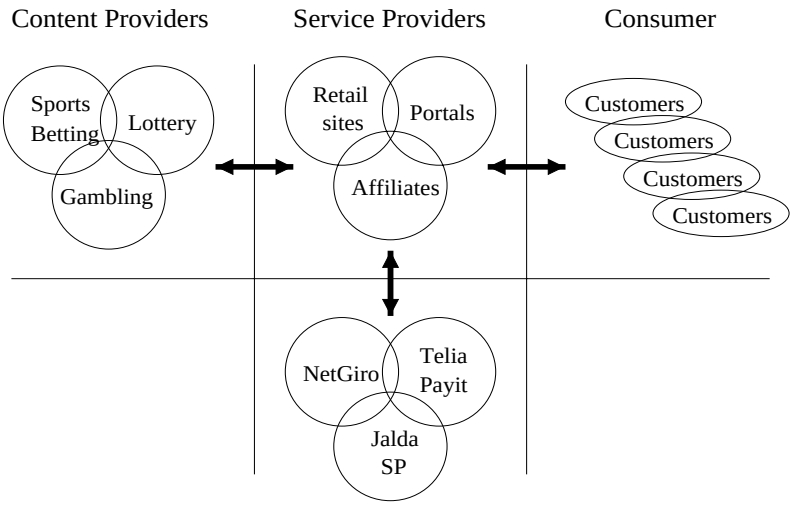

FIGURE 3. Gambling service

of paying the bill. If he wins money, however, the positive account balance can be used to buy services from other services providers associated with the same payment provider.

Identity theft. A stolen user identity and authentication token may be used to gamble on the account of another customer. This could be caused by bad or broken authentication mechanisms or customer negligance. In the intrusion detection area, this type of attack is often called *masquerading*.

The fraud cases for the *service provider* include:

Withholding revenue. The service provider may sell services for which he is withholding revenue. This could be the result of defects in the gambling software, preventing the content provider from verifying each transaction. A loss may be reported as usual, while a win is withheld from the content provider. The fact that the service provider acts as a clearinghouse increases the risk for fraud.

The fraud cases for the *payment provider* include:

Overcharging customer. The payment provider may claim payment for services not consumed by the customer. For transactions of small values, the customer may not notice the charges at first glance (salami attack).

The *content provider* fraud types include:

Unfair gambling odds. The content provider may manipulate the gambling application such that the odds of winning decrease. This is not only unfair to end customers but also to the service and payment providers that put their reputation on stake by offering the gambling service to their customers.

4.2 Fraud study

Five possible fraud scenarios were suggested for the gambling service. Here we study them using the proposed fraud model. Table 1 shows a classification of frauds based on the characteristics of the first two levels of the fraud model. The rest of the characteristics of each fraud are described below the table.

Fraud goal	**Actor**	**Fraud**
Evasion of payment	Customer	Subscription fraud
	SP	
	CP	
	EPP	
	ANP	
	NP	
	Non-customer	Identity theft

TABLE 1. Fraud classification

SLA violations	Customer	
	SP	Withholding revenue
	CP	Unfair gambling odds
	EPP	Overcharging customer
	ANP	
	NP	
	Non-customer	
Improper use	Customer	
	SP	
	CP	
	EPP	
	ANP	
	NP	
	Non-customer	
Sabotage	Customer	
	SP	
	CP	
	EPP	
	ANP	
	NP	
	Non-customer	

TABLE 1. Fraud classification

4.2.1 Subscription fraud

Fraud goal: Evasion of payment.

Actors involved: The customer commits fraud against the service provider and/or e-payment provider.

Service specific risks: It is desirable to gamble with no risk of losing. Service is not prepaid. It is difficult to identify a customer over a computer network.

Fraud technique: Social engineering fraud. Customer registration process is attacked.

4.2.2 Identity theft

Fraud goal: Evasion of payment.

Actors involved: The non-customer commits fraud against the customer and indirectly also against the service provider and/or the e-payment provider.

Service specific risks: It is desirable to gamble using someone else's money. It is possible to use another persons identity and authentication token.

Fraud technique: May be done either by social engineering or technical methods. Social engineering techniques include e.g. physically stealing an authentication token and "shoulder surfing". Technical methods include e.g. eavesdropping on network traffic to pick up authentication information.

4.2.3 Withholding revenue

Fraud goal: SLA violations.

Actors involved: The service provider commits fraud against the customer and indirectly also the content provider. What looks like a fraud where the service provider is withholding revenue may also be caused by an external fraudster who is in control of service components.

Service specific risks: The service provider is in a trusted position, as he mediates information about winnings and losses from the content provider to the customer. The customer may not be able to verify the transactions if he is not in close contact with the content provider. The content provider's software may have inadequate or defective monitoring functions.

Fraud technique: This may be a technical fraud if the service provider exploits a defect in the gambling software which prevents the content provider from verifying each transaction. It may be a non-technical fraud if the service provider exploits the content provider's poor fraud awareness and service management.

4.2.4 Over-charging customer

Fraud goal: SLA violations.

Actors involved: The e-payment provider (or an external fraudster) commits fraud against the customer and indirectly against the service provider.

Service specific risks: The payment provider is in a trusted position and can exploit trust. It can be difficult for the customer to verify the correctness of the transactions if he is not in close contact with the service provider or the customer may not be monitoring the transactions at all, which makes it a low risk fraud.

Fraud technique: Can be a non-technical fraud if the payment provider manually generates false information for the service provider. If the payment provider modifies billing software to generate a false billing log for the service provider, it may be considered a technical fraud.

4.2.5 Unfair gambling odds

Fraud goal: SLA violations.

Actors involved: The content provider (or an external fraudster) commits fraud against the customer and indirectly against the service provider.

Service specific risks: It is difficult for the customer and service provider to verify the quality of the service offered when the service content consists of remote use of software. The content provider can therefore "save" a reasonable amount of money at a low risk.

Fraud technique: Modify gambling software to generate fewer winners, lower winnings and/or unfair distribution of winnings. Log files with statistics may also be destroyed or modified to cover up the fraud.

4.3 Discussion

The fraud model may act as support when a new service is designed. By defining the actors, billing model, content and components, we can predict what known frauds are applicable in the new service. A complete database of known frauds with descriptions following the model would be of great help in this task. If it is found that some known problematic frauds are expected, a redesign of the service can be considered at an early stage to decrease the fraud risk.

The model may also be used to predict new frauds. By studying the combinations of fraud goals and actors in Table 1, we can consider for each combination whether the service encourages a fraud of this type and define the possible service risks. If we instead study e.g. combinations of actors and service risks, we can predict possible or probable fraud methods.

The goal of the work of defining possible frauds for a service is of course to develop "detection rules", or fraud indicators that can be used in the detection process. For example, a possible combination of fraud indicators to detect a submission fraud may be to look for new subscribers with low credit and high service usage.

Fraud indicators are important because they are needed regardless of the detection methods used. We have not tried to develop fraud indicators for the gambling service. To be able to derive usable fraud indicators from fraud characteristics, we must have a methodology. Hopefully, our fraud model will be of use in this area as well. This is an interesting research area that we intend to include in future work.

5. Detection techniques

Given the fraud model in the previous section, we will here discuss various techniques that can be applied in a new fraud management system and in particular address techniques already known from the area of intrusion detection. We will also investigate the possible benefit of combined ID-FD approaches.

5.1 Telecommunication Fraud Management Systems

5.1.1 Traditional FMS

A Telecommunication Fraud Management System (FMS) is an automated tool for detecting and managing frauds in telecommunication services. The FMS usually has an advanced operator interface and includes tools for manual fraud investigation. Most commercial Fraud Management Systems (FMS) use *Call Detail Records* (CDR) as input. CDRs are the basis for customer billing. These records can be used for creating customer profiles and for different kinds of threshold detection etc. It is also possible to use other sources of input data, such as SS7 signalling data [13].

Typically, call detail records (CDRs) are collected that contain information about the subscriber and the duration and destination etc. of a telephone call. A combination of thresholds, pattern recognition and statistical subscriber profiling can be used to find indications of fraud in the CDRs. Indications of fraud may also be found using very simple rules such as a list of hot numbers that are known to have been common in previous fraud investigations. Other rules trigger alarms when a certain parameter deviates substantially

from a given subscriber profile (e.g. the average duration of calls, the total number of calls during a certain period of time etc.)

5.1.2 Fraud detection

Fraud detection performed in services such as the one described in section 4.1 may have much in common with general computer system intrusion detection but there is a difference in the meaning of the terms. Fraud detection means detecting people misusing a service at the expense of some other party (e.g. a service provider). A service provider may lose money directly or indirectly because of fraud. ***Direct losses*** occur when resources are consumed for which the service provider does not receive payment. ***Indirect loss*** can occur if a user succeeds in damaging the reputation or market value of the service by for example denial-of-service attacks or by exploiting a service at the expense of another customer. If the service provider is not willing to take the cost, the company will lose goodwill and, eventually, customers.

Fraud detection includes social engineering scenarios but possibly excludes other interesting events that could be an indication of future attacks or fraud. Snooping around in the system may not be considered a fraud even though the person has violated the security policy.

5.1.3 FMS for IP based services

It is clear that the fraud management systems available today can not easily be applied to the new IP based services, mainly because no CDRs are generated by the new services. A FMS must thus be adapted to new types of input in order to be useful and/or some new type of CDR must be defined. What this new input data will look like is not yet clear, but ongoing initiatives are addressing the problem [12].

Billing data are often the main source of information in traditional fraud management systems. These will probably be the most useful data source in the new services as well. The main fraud detection system should therefore primarily detect potential frauds in the billing data. Much of the traditional fraud detection techniques can probably be used here. If the format of the billing data is standardized, a detection component can be general for all kinds of different services. However, billing data is not sufficient to make a complete detection system. One idea is to use separate detection components to monitor important service components that are not fully covered by the main detection system. Other forms of input data may be used here, e.g. network traffic and operating system log files. Furthermore, service components that are known to be vulnerable, should be subject to extra monitoring.

Since the new IP based services are predicted to be more dynamic, the detection system must also be dynamic. It should be possible to remove or add detection components to the system without a great deal of trouble. Correlation functions should also be easy to add for new components.

5.2 Intrusion Detection Systems

Intrusion detection can be defined as a technique to detect events in a system (or a domain) that violates a security policy. The security policy can be explicit or implicit. The term intrusion detection normally excludes detection of frauds committed without misusing some technical component of the system. Purely social engineering attacks, such as giving a false identity at registration for a user account, is difficult to detect by technical means, and is thus not likely to be considered by an intrusion detection system.

A great number of commercial Intrusion Detection Systems (IDS) have appeared in the last few years [1]. Although research has been going on in this area for at least two decades, only very simple techniques have been adapted in commercial systems. In research, however, all kinds of techniques are discussed. Most of the techniques used for fraud detection have also been used for intrusion detection. Most commercial intrusion detection systems use network packets and host-based log files, often from the operating system, as input. In some cases, application log files such as firewall logs or logs from a web server can be input for analysis.

The most common detection strategy in commercial IDSs is to create rules for known intrusive behavior. This strategy is called misuse detection and is often implemented as rule-based expert systems or as simple pattern matching. Events that these systems detect are e.g. network denial of service attacks like syn flooding and buffer overflow attacks that are used to gain increased privileges. The other detection strategy, that is used in only a few commercial systems, is anomaly detection. Anomaly detection means creating behavior profiles and then detect deviations from these profiles. It can be implemented using e.g. statistical methods or neural networks. This strategy can be used for detecting masqueraders, password guessing, etc. In [8], a discussion of detection strategies and implementation methods can be found. Events are in most cases not correlated to any higher degree in intrusion detection systems, which means that some attacks are missed and some causes floods of alarms.

5.3 Fraud detection vs intrusion detection - differences and similarities

A detection system in a complex datacom service may include both fraud and intrusion detection to cover all events of interest. There are some fundamental differences in the design criteria between systems for intrusion and fraud detection. In general, a fraud detection system is designed to detect fraud above a certain threshold. That is, the detection threshold can be adjusted such that insignificant cases of fraud do not generate floods of alarms. An intrusion detection system is designed to detect intrusions and intrusion attempts. There is no such thing as an insignificant intrusion, which makes it difficult to tune an intrusion detection system not to raise too many false alarms. IDS techniques are not directly applicable to fraud detection. Fraud detection can be seen as an application specific form of intrusion detection. Application specific input and application specific rules for what is considered “misuse” must be applied. It would be an advantage if it was possible to adapt the IDS in such a way that it could be used for fraud detection. However, most intrusion detection systems of today can not be adapted in a straightforward and effective manner.

A combination of IDS and FMS may be a way of achieving better coverage of the possible frauds and intrusions in future services. One key advantage of utilizing fraud detection techniques is that it enables an organization to put a price tag on losses caused by fraud. This can not easily be done for intrusion detection as it is difficult to predict the economic consequences caused by an intrusion. This could be used within an organization to help motivate security improving investments.

5.3.1 Vulnerability detection

There are also differences in the detection strategies in FMS and IDS. One possible strategy is to have a list of all vulnerabilities in the system and try to detect the use of these vulnerabilities. This strategy is often called misuse detection in the intrusion detection area. Here the focus is on detecting the technique used by the attacker or fraudster. This is the major method used in commercial intrusion detection systems today. The list of known intrusion/fraud techniques often becomes long and tiresome to maintain. It is also easy to leave out important entries and to miss variations of attacks. Currently unknown techniques will also be excluded. To collect and analyze the data necessary to detect the use of all different vulnerabilities is resource demanding.

5.3.2 Consequence detection

A detection method entirely based on these vulnerability lists is probably not feasible in a complex system such as the one we are dealing with. To reduce complexity, consequence detection may be used instead. This means looking at the “attack goal” rather than the method the attacker has used. In the fraud case, this often means detecting that less money is flowing in to the service than expected, or detecting a more extensive usage of the service than expected.

The attack method is of course interesting, but it can be investigated in more detail after a fraud or intrusion has been detected. The detection strategy in FMS often resembles consequence detection. The consequence of a fraud may for example be that a user has access to a service he is not paying for or that a denial-of-service attack is in progress. This method may require fewer auditing points and less resources.

Among the disadvantages of consequence detection are that attempted attacks can not be detected and that the detection is not always timely. It is interesting to study attempted attacks because they give a hint of how large the threat is and what methods are most commonly used. The timeliness of the detection may be important because some attacks can cause damage before they are detected. The timeliness problems of this method are caused by the fact that the intrusion is not detected until the attacker performs an act considered to be the “goal” of the attack. For example, he may have been snooping around and placed Trojans in the system before he completes the attack. He also may have modified or removed log data.

The disadvantages of consequence detection may be reduced or remedied by combining it with detection targeted at known vulnerabilities. Especially, extra effort may be used to monitor components of the system that are considered sensitive.

5.4 Combining fraud and intrusion detection

Intrusion and fraud detection applications serve different purposes even though they share many characteristics and functions. This section highlights issues that need to be addressed when combining fraud and intrusion detection.

To combine FMS and IDS in future fraud detection systems, we must have an idea about how to do this effectively. As we discussed in the previous chapter, the typical detection strategies in the two types of systems can complement each other. However, a detection system consists of more than a detection strategy and these other components also needs to be discussed.

Figure 4 shows the components that make up a typical IDS/FMS system.

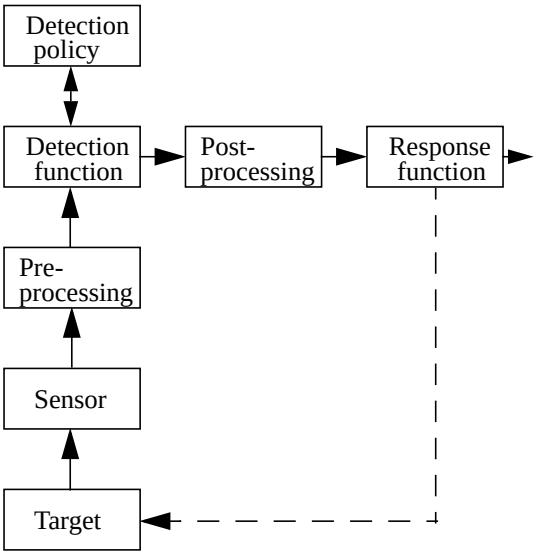

FIGURE 4. Components of an IDS/FMS system

Targets and sensors. The targets that are monitored are usually some kind of network element or computer system that produces the audit data to be analyzed. The type of analysis determines the features that are of interest, such as IP header data, performance and/or billing data. For traditional telecom fraud detection systems, CDRs are analyzed for detecting fraudulent behavior. However, as future telecom services migrate towards the IP protocol suite, the differences between IDS and FMS data collection will be less noticeable.

Preprocessing. Preprocessing involves filtering and formatting audit data such that it can be applied to different analysis (detection) mechanisms. Our studies of IDS/FMS lead us to believe that the preprocessing phase is similar for both types of systems.

Detection function and policy. Simple rule based detection schemes are successfully being used for both types of systems in current products and prototypes. Pattern matching and threshold detection mechanisms are examples of rule based mechanisms. In addition, Telecom FMS often build and manage customer profiles describing some characteristics of a subscriber/user. For example, a profile may be created that contains the average length of a telephone call over a certain period of time. Similarly, anomaly based IDSs seek to find anomalies over some parameter of interest. Our conclusion is that the detection function is not limited to either IDS or FMS but could be used for both applications.

Postprocessing. The postprocessing of alarms can involve several different actions, such as classification, prioritization and storage of alarms. In FMS, a case building facility is often utilized to cluster events that relate to each other in some respect (case building). This is a feature that may also be useful for IDS. However, few existing IDS products have this functionality. In most cases, if an attack triggers multiple alarms, it is up to the operator to interpret these as a single event (case).

Response. The response function is perhaps the component that differs the most between IDS and FMS. Once an intrusion has been detected by an IDS, automatic response actions can be performed that terminate the attack or try to limit the damage caused by an attack. Automatic response functions must be used with caution, however. It is always a balance between the value of the assets that the IDS tries to protect and the availability of the service that the asset realise. FMS typically report indications of fraud that can not necessarily be stopped by automatic response actions. Indications of fraud will in most cases require an extensive investigation involving several non-technical parties such as the company financial department, law enforcement agencies etc.

Architectural issues. To make the detection system effective, alarms from different detection components should be managed centrally by the combined FD/ID system.

Other functions can also be managed centrally, such as storage of log data. The components must therefore be compatible or use a common information exchange format. Correlation of data from different data sources and different detection components is also desirable in a system like this and it will require far-reaching standardization in order to be effective.

More issues are likely to appear in efforts to combine fraud and intrusion detection in practice. It seems as though the combination of knowledge from the two areas can result in a much better system, even though many details of how to accomplish it still remain.

6. Conclusions

Although the areas of fraud detection and intrusion detection have many common denominators, we found that there is very little crossover research. It becomes clear that integration of the two areas would be beneficial in looking at the new problems that will appear in new telecom services, especially in IP based services. We also found that there is a lack of applicable fraud models, and we believe that our fraud model may be of use in predicting frauds for new services and for the design of reusable modules for fraud detection systems. However, there is still some work to be done on

the model, and we have plan to test it in the design of a fraud detection system for an IP based service.

7. References

[1] D. W. Abbott, I. P. Matkovsky, and J. F. Elder. An evaluation of high-end data mining tools for fraud detection. In *1998 IEEE International Conference on Systems, Man, and Cybernetics*, volume 3, pages 2836-2841. IEEE, October 11-14 1998. ISBN: 0-7893-4778-1.

[2] F. Bonchi, F. Giannotti, G. Mainetto and D. Pedreschi. A classification-based methodology for planning audit strategies in fraud detection. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 175-184, August 1999.

[3] Michael Collins. Telecommunications Crime - Part 1. In *Computers & Security, 18 (1999)*, pages 577-586

[4] Michael Collins. Telecommunications Crime - Part 2. In *Computers & Security, 18 (1999)*, pages 683-692

[5] Michael Collins. Telecommunications Crime - Part 3. In *Computers & Security, 19 (2000)*, pages 141-148

[6] Hervé Debar, Marc Dacier, and Andreas Wespi. Towards a Taxonomy of Intrusion-Detection Systems. In *Computer Networks*, volume 31 (issue 8), pages 805-822, April 1999.

[7] B. Garner and F. Chen. Hypothesis generation paradigm for fraud detection. In *Proceedings of TENCON '94, IEEE Region 10's Ninth Annual International Conference*. IEEE, 1994. Theme: Frontiers of Computer Technology.

[8] Lawrence R. Halme, R. Kenneth Bauer. AINT misbehaving - a taxonomy of anti-intrusion techniques. In *Proceedings of the 18th National Information Systems Security Conference*, pages 163-172, October 1995, Baltimore, MD, USA. Published by National Institute of Standards and Technology/National Computer Security Center.

[9] Paul Helman and Gunar Liepins. Statistical foundations of audit trail analysis for the detection of computer misuse. In *IEEE Transactions on Software Engineering*, vol 19 (issue 9), pages 886-901, September 1993.

[10] Peter Hoath. What's new in telecoms fraud? *Computer Fraud & Security*, volume 16 (issue 2):13-19, February 1999. Elsevier Science Ltd.

[11] Jaakko Hollmén, Volker Tresp, and Olli Simula. A self-organizing map for clustering probabilistic models. In *Ninth International Conference on Artificial Neural Networks* (ICANN 99), volume 2, pages 946-951. IEEE, September 7-10 1999.

[12] IPDR Working group. Network Data Management - Usage (NDM-U) for IP-based services, www.ipdr.org

[13] ISO SS7 standard

[14] Håkan Kvarnström. A survey of commercial tools for intrusion detection. Technical Report TR99-8, Chalmers University of Technology.

[15] Ulf Lindqvist and Erland Jonsson. How to systematically classify computer security intrusions. In *Proceedings of the 1997 Symposium on Security and Privacy*, pages 154-163, Oakland, California, May 4-7, 1997.