Statistical Fraud Detection: A Review

Richard J. Bolton and David J. Hand

January 2002

Abstract:

Fraud is increasing dramatically with the expansion of modern technology and the global superhighways of communication, resulting in the loss of billions of dollars worldwide each year. Although prevention technologies are the best way of reducing fraud, fraudsters are adaptive and, given time, will usually find ways to circumvent such measures. Methodologies for the detection of fraud are essential if we are to catch fraudsters once fraud prevention has failed. Statistics and machine learning provide effective technologies for fraud detection and have been applied successfully to detect activities such as money laundering, e-commerce credit card fraud, telecommunication fraud, and computer intrusion, to name but a few. We describe the tools available for statistical fraud detection and the areas in which fraud detection technologies are most used.

Keywords:

Fraud detection, fraud prevention, statistics, machine learning, money laundering, computer intrusion, e-commerce, credit cards, telecommunications.

Author’s note: Richard J. Bolton is Research Associate and David J. Hand Professor of Statistics, Department of Mathematics, Imperial College, 180 Queen’s Gate, London SW7 2BZ, UK. Contact email: {r.bolton, d.j.hand @ic.ac.uk}

1. Introduction

The Concise Oxford Dictionary defines fraud as ‘*criminal deception; the use of false representations to gain an unjust advantage*'. Fraud is as old as humanity itself, and can take an unlimited variety of different forms. However, in recent years, the development of new technologies (which have made it easier for us to communicate and helped increase our spending power) has also provided yet further ways in which criminals may commit fraud. Traditional forms of fraudulent behaviour such as money laundering have become easier to perpetrate and have been joined by new kinds of fraud such as mobile telecommunications fraud and computer intrusion.

We begin by distinguishing between fraud prevention and fraud detection. Fraud *prevention* describes measures to stop fraud occurring in the first place. These include elaborate designs, fluorescent fibres, multitone drawings, watermarks, laminated metal strips, and holographs on banknotes, PINs for bankcards, Internet security systems for credit card transactions, SIM cards for mobile phones, and passwords on computer systems and telephone bank accounts. Of course, none of these methods are perfect, and in general, a compromise has to be struck between expense and inconvenience (for example, to a customer), on the one hand, and effectiveness on the other.

In contrast, fraud *detection* involves identifying fraud as quickly as possible once it has been perpetrated. Fraud detection comes into play once fraud prevention has failed. In practice, of course fraud detection must be used continuously, as one will typically be unaware that fraud prevention has failed. We can try to prevent credit card fraud by guarding our cards assiduously, but if nevertheless the card’s details are stolen, then we need to be able to detect, as soon as possible, that fraud is being perpetrated.

Fraud detection is a continuously evolving discipline. Whenever it becomes known that one detection method is in place, criminals will adapt their strategies and try others. Of course, new criminals are also constantly entering the field. Many of these will not be aware of the fraud detection methods which have been successful in the past, and will adopt strategies which lead to identifiable frauds.

This means that the earlier detection tools need to be applied as well as the latest developments.

The development of new fraud detection methods is made more difficult by the fact that the exchange of ideas in fraud detection is severely limited. It does not make sense to describe fraud detection techniques in great detail in the public domain, as this gives criminals the information that they require in order to evade detection. Data sets are not made available and results are often censored, making them difficult to assess (for example, Leonard 1993).

Many fraud detection problems involve huge data sets that are constantly evolving. For example, the credit card company Barclaycard carries approximately 350 million transactions a year in the UK alone (Hand *et al*, 2000), The Royal Bank of Scotland, which has the largest credit card merchant acquiring business in Europe, carries over a billion transactions a year, and AT&T carries around 275 million calls each weekday (Cortes and Pregibon, 1998). Processing these in a search for fraudulent transactions or calls requires more than mere novelty of statistical model, and also needs fast and efficient algorithms: data mining techniques are relevant. These numbers also indicate the potential value of fraud detection: if 0.1% of a 100 million transactions are fraudulent, each losing the company just £10, then overall the company loses £1 million.

Statistical tools for fraud detection are many and varied, since data from different applications can be diverse in both size and type, but there are common themes. Such tools are essentially based on comparing the observed data with expected values, but expected values can be derived in various ways, depending upon the context. They may be single numerical summaries of some aspect of behaviour, they are often simple graphical summaries in which an anomaly is readily apparent, but they are also often more complex (multivariate) behaviour profiles. Such behaviour profiles may be based on past behaviour of the system being studied (for example, the way a bank account has been previously used), or by extrapolation from other, similar, systems. Things are often further complicated by the fact that, in some domains (e.g. trading on the stock market) a given actor may behave in a fraudulent manner some of the time and not at other times.

Statistical fraud detection methods may be ‘supervised’ or ‘unsupervised’. In supervised methods, samples of both fraudulent and non-fraudulent records are used to construct models which allow one to assign new observations into one of the two classes. Of course, this requires one to be confident about the true classes of the original data used to build the models. It also requires that one has examples of both classes. Furthermore, it can only be used to detect frauds of a type which have previously occurred.

In contrast, unsupervised methods simply seek those accounts, customers, etc. which are most dissimilar from the norm. These can then be examined more closely. Outliers are a basic form of non-standard observation. Tools used for checking data quality can be used, but the detection of accidental errors is a rather different problem from the detection of deliberately falsified data or data which accurately describe a fraudulent pattern.

This leads us to note the fundamental point that we can seldom be certain, by statistical analysis alone, that a fraud has been perpetrated. Rather, the analysis should be regarded as alerting us to the fact that an observation is anomalous, or more likely to be fraudulent than others – so that it can then be investigated in more detail. One can think of the objective of the statistical analysis as being to return a *suspicion score* (where we will regard a higher score as more suspicious than a lower one). The higher the score is, then the more unusual is the observation, or the more like previously fraudulent values it is. The fact that there are many different ways in which fraud can be perpetrated, and many different scenarios in which it can occur, means that there are many different ways of computing suspicion scores.

Suspicion scores can be computed for each record in the database (for each customer with a bank account or credit card, for each owner of a mobile phone, for each desktop computer, and so on), and these can be updated as time progresses. These can then be rank ordered, and investigative attention can be focussed on those with the highest scores, or on those which exhibit a sudden increase. Here issues of cost enter: given that it is too expensive to undertake a detailed investigation of all records, one concentrates investigation on those thought to be most likely to be fraudulent.

One of the difficulties with fraud detection is that typically there are many legitimate records for each fraudulent one. A detection method which correctly identifies 99% of the legitimate records as legitimate and 99% of the fraudulent records as fraudulent might be regarded as a highly effective system. However, if only one in a thousand records are fraudulent, then, on average in every 100 that the system flags as fraudulent, only about 9 will in fact be so. In particular, this means that to identify those 9 requires detailed examination of all 100 – at possibly considerable cost. This leads us to a more general point: fraud can be reduced to as low a level as one likes, but only by virtue of a corresponding level of effort and cost. In practice, some compromise has to be reached, often a commercial compromise, between the cost of detecting a fraud and the savings to be made by detecting it. Sometimes the issues are complicated by, for example, the adverse publicity accompanying fraud detection. At a business level, revealing that a bank is a significant target for fraud, even if much has been detected, does little to inspire confidence, and at a personal level, taking action which implies to an innocent customer that they may be suspected of fraud is obviously detrimental to good customer relations.

The body of this paper is structured according to different areas of fraud detection. Clearly we cannot hope to cover all areas in which statistical methods can be applied. Instead, we have selected a few areas where such methods are used, and where there is a body of expertise and of literature describing them. However, before looking at the details of different application areas, Section 2 provides a brief overview of some tools for fraud detection.

2. Fraud detection tools

As we mentioned above, fraud detection can be ‘supervised’ or ‘unsupervised’. Supervised methods use a database of known fraudulent/legitimate cases from which to construct a model which yields a suspicion score for new cases. Traditional statistical classification methods (Hand, 1981; McLachlan, 1992) such as linear discriminant analysis and logistic discrimination have proved to be effective tools for many applications, but more powerful tools (Ripley, 1996;

Hand, 1997; Webb, 1999), especially neural networks, have also been extensively applied. Rule-based methods are supervised learning algorithms that produce classifiers using rules of the form: *If* {certain conditions}, *Then* {a consequent}. Examples of such algorithms include BAYES (Clark and Niblett, 1989), FOIL (Quinlan 1990) and RIPPER (Cohen 1995). Tree-based algorithms such as CART (Brieman *et al*, 1984) and C4.5 (Quinlan 1993) produce classifiers of a similar form. Combinations of some or all of these algorithms can be combined using meta-learning algorithms to improve prediction in fraud detection e.g. Chan *et al* (1999).

Major considerations when building a supervised tool for fraud detection include those of uneven class sizes and different costs of different types of misclassification. We must also take into consideration the costs of investigating observations and the benefits of identifying fraud. Moreover, often class membership is uncertain. For example, credit transactions may be labelled incorrectly; a fraudulent transaction may remain unobserved and thus labelled as legitimate (and the extent of this may remain unknown), or a legitimate transaction may be misreported as being fraudulent. Some work has addressed misclassification of training samples (for example, Lachenbruch (1966, 1974); Chhikara and McKeon (1984)) but not in the context of fraud detection as far as we are aware. Issues such as these are discussed by Chan and Stolfo (1998) and Provost and Fawcett (2001).

Link analysis relates known fraudsters to other individuals, using record linkage and social network methods (Wasserman and Faust, 1994). For example, in telecommunications networks, security investigators have found that fraudsters seldom work in isolation from each other; also, after an account has been disconnected for fraud, the fraudster will often call the same numbers from another account (Cortes *et al*, 2001). Telephone calls from an account can thus be linked to fraudulent accounts to indicate intrusion. A similar approach has been taken in money laundering (Goldberg and Senator, 1995; Senator *et al*, 1995; Goldberg and Senator, 1998).

Unsupervised methods are used when there are no prior sets of legitimate and fraudulent observations. Techniques employed here are usually a combination of profiling and outlier detection methods. We model a baseline distribution that represents normal behaviour and then attempt to detect observations that show greatest departure from this norm. There are similarities to author identification in text analysis. Digit analysis using Benford's law is an example of such a method. Benford's law (Hill 1996) says that the distribution of the first significant digits of numbers drawn from a wide variety of random distributions will have (asymptotically) a certain form. Until recently, this law was regarded as merely a mathematical curiosity with no apparent useful application. However, Nigrini and Mittermaier (1997) and Nigrini (1999) show that Benford’s law can be used to detect fraud in accounting data. The premise behind fraud detection using tools such as Benford’s law is that fabricating data which conform to Benford's law is difficult.

Fraudsters adapt to new prevention and detection measures, so fraud detection needs to be adaptive and evolve over time. However, legitimate account users may gradually change their behaviour over a longer period of time, and it is important to avoid spurious alarms. Models can be updated at fixed time points or continuously over time - see, for example Burge and Shawe-Taylor (1997), Fawcett and Provost (1997a), Cortes *et al* (2001) and Senator (2000).

Although the basic statistical models for fraud detection can be categorised as supervised or unsupervised, the application areas of fraud detection cannot be described so conveniently. Their diversity is reflected in their particular operational characteristics and the variety and quantity of data available, both features that drive the choice of a suitable fraud detection tool.

3. Credit Card Fraud

The extent of credit card fraud is difficult to quantify, partly because companies are often loath to release fraud figures in case they frighten the spending public, and partly because the figures change (probably grow) over time. Various estimates have been given. For example, Leonard (1993) suggested the cost of Visa/Mastercard fraud in Canada in 1989, 1990, and 1991 was 19, 29, and 46 million Canadian dollars, respectively. Ghosh and Reilly (1994) suggest a figure of 850 million US dollars per year for all types of credit card fraud in the US, and Aleskerov *et al* (1997) cite estimates of $700 million in the US each year for Visa/Mastercard, and $10 billion worldwide in 1996. Microsoft's Expedia set aside $6 million for credit card fraud in 1999 (Patient, 2000). Total losses through credit card fraud in the UK have been growing rapidly over the last four years (1997 - £122m; 1998 – £135m; 1999 - £188m; 2000 - £293m. Source: Association for Payment Clearing Services, London (APACS)) and recently APACS has reported £373.7m losses in the 12 months to end August 2001. Jenkins (2000) says 'for every £100 you spend on a card in the UK, 13p is lost to fraudsters.' Matters are complicated by issues of exactly what one includes in the fraud figures. For example, 'bankruptcy fraud' arises when the cardholder makes purchases for which he/she has no intention of paying and then files for personal bankruptcy, leaving the bank to cover the losses. Since these are generally regarded as charge-off losses, they often are not included in fraud figures.

However, they can be substantial: Ghosh and Reilly (1994) cite one estimate of $2.65 billion for bankruptcy fraud in 1992.

It is in a company and card issuer's interests to prevent fraud or, failing this, to detect fraud as soon as possible. Otherwise consumer trust in both the card and company decreases and revenue is lost, in addition to the direct losses made through fraudulent sales. Because of the potential for loss of sales due to loss of confidence, in general the merchants assume responsibility for fraud losses, even when the vendor has obtained authorisation from the card issuer.

Credit card fraud may be perpetrated in various ways (a description of the credit card industry and how it functions is given in Blunt and Hand, 2000), including simple theft, application fraud, and counterfeit cards. In all of these, the fraudster uses a physical card, but physical possession is not essential in order to perpetrate credit card fraud: one of the major fraud areas is 'cardholder-not-present' fraud, where only the card's details are given (for example, over the phone).

Use of a stolen card is perhaps the most straightforward type of credit card fraud. In this case, the fraudster typically spends as much as possible in as short a space of time as possible, before the theft is detected and the card stopped, so that detecting the theft early can prevent large losses.

Application fraud arises when individuals obtain new credit cards from issuing companies using false personal information. Traditional credit scorecards (Hand and Henley, 1997) are used to detect customers who are likely to default, and the reasons for this may include fraud. Such scorecards are based on the details given on the application forms, and perhaps also on other details, such as bureau information. Statistical models which monitor behaviour over time can be used to detect cards which have been obtained from a fraudulent application (e.g. a first time card holder who runs out and rapidly makes many purchases should arouse suspicion). With application fraud, however, urgency is not so important to the fraudster, and it might not be until accounts are sent out or repayment dates begin to pass that fraud is suspected.

Cardholder-not-present fraud occurs when the transaction is made remotely, so that only the card's details are needed, and a manual signature and card imprint are not required at the time of purchase. Such transactions include telephone sales and online transactions, and this type of fraud accounts for a high proportion of losses. To undertake such fraud it is necessary to obtain the details of the card without the cardholder's knowledge. This is done in various ways, including 'skimming', where employees illegally copy the magnetic stripe on a credit card by swiping it through a small handheld card reader, 'shoulder surfers' who enter card details into a mobile phone while standing behind a purchaser in a queue, and people posing as credit card company employees taking details of credit card transactions from companies over the phone. Counterfeit cards, currently the largest source of credit card fraud in the UK (source: APACS), can also be created using this information. Transactions made by fraudsters using counterfeit cards and making cardholder-not-present purchases can be detected through methods which seek changes in transaction patterns, as well as checking for particular patterns which are known to be indicative of counterfeit.

Credit card databases contain information on each transaction. This information includes such things as merchant code, account number, type of credit card, type of purchase, client name, size of transaction, date of transaction, etc. Some of these are numerical (e.g. transaction size) and others are nominal categorical (e.g. merchant code, which can have hundreds of thousands of categories) or symbolic. The mixed data types has led to a wide variety of statistical, machine learning, and data mining tools being applied.

Suspicion scores to detect whether an account has been compromised can be based on models of individual customers' previous usage patterns, standard expected usage patterns, particular patterns which are known to be often associated with fraud, and also on supervised models. A simple example of the patterns exhibited by individual customers is given in Figure 16 of Hand and Blunt (2001), which shows how the slopes of cumulative credit card spend over time are remarkably linear. Sudden jumps in these curves, or sudden changes of slope (transaction or expenditure rate suddenly exceeding some threshold), merit investigation. Likewise, some customers practice ‘jam jarring’ – restricting particular cards to particular types of purchases (e.g. using a given card for petrol purchases only, and a different one for supermarket purchases), so that usage of a card to make an unusual type of purchase can trigger an alarm for such customers. At a more general level, suspicion scores can also be based on expected overall usage profiles. For example, first time credit card users are typically initially fairly tentative in their usage, whereas those transferring loans from another card are generally not so reticent. Finally, examples of overall transaction patterns known to be intrinsically suspicious are the sudden purchase of many small electrical items or jewelry (goods which permit easy black market resale) and the immediate use of a new card in a wide range of different locations.

We commented above that, for obvious reasons, there is a dearth of published literature on fraud detection. Of that which has been published, much of it appears in the methodological data analytic literature, where the aim is to illustrate new data analytic tools by applying them to the detection of fraud, rather than being to describe methods of fraud detection per se. Furthermore, since anomaly detection methods are very context dependent, much of the published literature in the area concentrates on supervised classification methods. In particular, rule-based systems and neural networks have attracted interest. Researchers who have used neural networks for credit card fraud detection include Ghosh and Reilly (1994), Aleskerov *et al* (1997), Dorronsoro *et al* (1997), and Brause *et al* (1999), mainly in the context of supervised classification. HNC Software has developed *Falcon*, a software package that relies heavily on neural network technology to detect credit card fraud.

Supervised methods, using samples from the fraudulent/non-fraudulent classes as the basis to construct classification rules detecting future cases of fraud, suffer from the problem of unbalanced class sizes mentioned above: the legitimate transactions generally far outnumber the fraudulent ones. Brause *et al* (1999) say that, in their database of credit card transactions, 'the probability of fraud is very low (0.2%) and has been lowered in a preprocessing step by a conventional fraud detecting system down to 0.1%.' Hassibi (2000) remarks 'Out of some 12 billion transactions made annually, approximately 10 million – or one out of every 1200 transactions – turn out to be fraudulent. Also, 0.04% (4 out of every 10,000) of all monthly active accounts are fraudulent.' It follows from this sort of figure that simple misclassification rate cannot be used as a performance measure: with a bad rate of 0.1%, simply classifying every transaction as legitimate will yield an error rate of only 0.001. Instead, one must either minimize an appropriate cost-weighted loss or fix some parameter (such as the number of cases one can afford to investigate in detail) and then try to maximise the number of fraudulent cases detected subject to this.

Stolfo, Fan *et al* (1997) and Stolfo, Prodromidis *et al* (1997) outline a meta-classifier system for detecting credit card fraud, based on the idea of using different local fraud detection tools within each different corporate environment, and merging the results to yield a more accurate global tool. This work is elaborated in Chan and Stolfo (1998), Chan *et al* (1999), and Stolfo *et al* (2000), who describe a more realistic cost model to accompany the different classification outcomes. Wheeler and Aitken (2000) have also explored the combination of multiple classification rules.

4. Money laundering

Money laundering is the process of obscuring the source, ownership, or use of funds, usually cash, that are the profits of illicit activity. The size of the problem is indicated in a 1995 US Office of Technology Assessment report (US Congress, 1995): 'Federal agencies estimate that as much as $300 billion is laundered annually, worldwide. From $40 billion to $80 billion of this may be drug profits made in the United States.' Prevention is attempted by means of legal constraints and requirements – the burden of which is gradually increasing – and there has been much debate recently about the use of encryption. However, no prevention strategy is foolproof, and detection is essential. In particular, the 11th September terrorist attacks on New York and the Pentagon have focused attention on the detection of money laundering in an attempt to starve terrorist networks of funds.

Wire transfers provide a natural domain for laundering: according to the OTA report, each day in 1995 about half a million wire transfers, valued at more than two trillion US dollars, were carried out using the Fedwire and CHIPS systems, along with almost a quarter of a million transfers using the SWIFT system. It is estimated that around 0.05% to 0.1% of these transactions were laundering transactions. Sophisticated statistical and other on-line data analytic procedures are needed to detect such laundering activity. Since it is now becoming a legal requirement to show that all reasonable means have been used to detect fraud, we may expect to see even greater application of such tools.

Wire transfers contain items such as date of transfer, identity of sender, routing number of originating bank, identity of recipient, routing number of recipient bank, and the amount transferred. Sometimes those fields not needed for transfer are left blank, free text fields may be completed in different ways, and, worse still but inevitable, sometimes the data has errors. Automatic error detection (and correction) software has been developed, based on semantic and syntactic constraints on possible content, but, of course, this can never be a complete solution. Things are also complicated by the fact that banks do not share their data. Of course, banks are not the only bodies that transfer money electronically, and other businesses have been established precisely for this purpose (the OTA report (US Congress, 1995) gives the estimated number of such businesses as 200,000).

The detection of money laundering presents difficulties not encountered in areas such as, for example, the credit card industry. Whereas credit card fraud comes to light fairly early on, in money laundering it may be years before individual transfers or accounts are definitively and legally identified as part of a laundering process. While, in principle (assuming records have been kept), one could go back and trace the relevant transactions, in practice not all of them would be identified, so detracting from their use in supervised detection methods. Furthermore, there is typically less extensive information available for the account holders in investment banks than there is in retail banking operations. Developing more detailed customer record systems might be a good way forward.

As with other areas of fraud, money laundering detection works hand in hand with prevention. In 1970, for example, in the US the Bank Secrecy Act required that banks report all currency transactions of over $10,000 to the authorities. However, also as in other areas of fraud, the perpetrators adapt their modus operandi to match the changing tactics of the authorities. So, following the requirement of banks to report currency transactions of over $10,000, the obvious strategy was developed of dividing larger sums into multiple amounts of less than $10,000, and depositing these in different banks (a practice termed *smurfing* or *structuring*). In the US, this is now illegal, but the way the money launderers adapt to the prevailing detection methods can lead one to the pessimistic perspective that only the incompetent money launderers are detected. This, clearly, also limits the value of supervised detection methods: the patterns detected will be those patterns which were characteristic of fraud in the past, but which may no longer be so. Other strategies used by money launderers, which limit the value of supervised methods include switching between wire and physical cash movements, the creation of shell businesses, false invoicing, and, of course, the fact that a single transfer, in itself, is unlikely to appear to be a laundering transaction. Furthermore, because of the large sums involved, money launderers are highly professional, often with contacts in the banks who can feed back details of the detection strategies being applied.

The number of currency transactions over $10,000 in value increased dramatically after the mid-1980s, to the extent that the number of reports filed is huge (over 10 million in 1994, with total worth of around $500 billion), and this in itself can cause difficulties. In an attempt to cope with this, the Financial Crimes Enforcement Network (FinCEN) of the US Department of the Treasury processes all such reports using the FinCEN Artificial Intelligence System (FAIS), described below. More generally, banks are also required to report any suspicious transactions, and about 0.5% of Currency Transaction Reports are so flagged.

Money laundering involves three steps:

- *placement*: the initial introduction of the cash into the banking system or legitimate business (for example, transferring the banknotes obtained from retail drugs transactions into a cashier's cheque). One way to do this is to pay vastly inflated amounts for goods imported across international frontiers. Pak and Zdanowicz (1994) describe statistical analysis of trade databases to detect anomalies in government trade data such as charging $1,694 a gram for imports of the drug erythromycin, compared with eight cents a gram for exports.

- *layering*: carrying out multiple transactions through multiple accounts with different owners at different financial institutions in the legitimate financial system;

- *integration*: merging the funds with money obtained from legitimate activities.

Detection strategies can be targeted at various levels. In general (and in common with some other areas in which fraud is perpetrated), it is very difficult or impossible to characterise an individual transaction as fraudulent. Rather transaction *patterns* must be identified as fraudulent or suspicious. A single deposit of just under $10,000 is not suspicious, but multiple such deposits are; a large sum being deposited is not suspicious, but a large sum being deposited and instantly withdrawn is. In fact, one can distinguish several levels of (potential) analysis: the individual transaction level, the account level, the business level (and, indeed, individuals may have multiple accounts), and the 'ring' of businesses level. Analyses can be targeted at particular levels, but more complex approaches can examine several levels simultaneously (there is an analogy here with speech recognition systems: simple systems focused at the individual phoneme and word levels are not as effective as those which try to recognise these elements in a higher level context of the way words are put together when used). In general, link analysis identifying groups of participants involved in transactions, plays a key role in most money laundering detection strategies. Senator *et al* (1995) say: 'Money laundering typically involves a multitude of transactions, perhaps by distinct individuals, into multiple accounts with different owners at different banks and other financial institutions. Detection of large-scale money laundering schemes requires the ability to reconstruct these patterns of transactions by linking potentially related transactions and then to distinguish the legitimate sets of transactions from the illegitimate ones. This technique of finding relationships between elements of information, called *link analysis*, is the primary analytic technique used in law enforcement intelligence (Andrews and Peterson, 1990).' An obvious and simplistic illustration is the fact that a transaction with a known criminal may rouse suspicion. More subtle methods are based on recognition of the sort of businesses that money laundering operations transact with. Of course, these are all supervised methods, and are subject to the weaknesses that those responsible may evolve their strategies. Similar tools are used in detecting telecoms fraud, as outlined in the following section.

Rule-based systems have been developed, often with the rules based on experience ('flag transactions from countries X and Y', 'flag accounts showing a large deposit followed immediately by a similar sized withdrawal'). Structuring can be detected by computing the cumulative sum of amounts entering an account over a short window, such as a day. Other methods have been developed based on straightforward descriptive statistics, such as rate of transactions, proportion of transactions which are suspicious, etc. The use of the Benford distribution is an extension of this idea. Although one may not usually be interested in detecting changes in an account's behaviour, methods such as peer group analysis (Bolton and Hand, 2001) and break detection (Goldberg and Senator, 1997), can be applied in detecting money laundering.

One of the most elaborate money laundering detection systems is the US Financial Crimes Enforcement Network AI system (FAIS), described in Senator *et al* (1995) and Goldberg and Senator (1998). This allows users to follow trails of linked transactions. It is built round a ‘blackboard’ architecture, in which program modules can read and write to a central database containing details of transactions, subjects, and accounts. A key component of the system is its suspicion score. This is a rule-based system based on an earlier system developed by the US Customs Service in the mid-1980s. The system computes suspicion scores for various different types of transaction and activity. Simple Bayesian updating is used to combine evidence suggesting that a transaction or activity is illicit, to yield an overall suspicion score. Senator *et al* (1995) includes a brief but interesting discussion of an investigation of whether case-based reasoning (c.f. nearest neighbour methods) and classification tree techniques could usefully be added to the system.

The American National Association of Securities Dealers, Inc. uses an *Advanced Detection System* (ADS) (Kirkland *et al*, 1999; Senator, 2000) to flag ‘patterns or practices of regulatory concern’. ADS uses a rule pattern matcher and a time-sequence pattern matcher, and (like FAIS) places great emphasis on visualisation tools. Also as with FAIS, data mining techniques are used to identify new patterns of potential interest.

A different approach to detecting similar fraudulent behaviour is taken by SearchSpace Ltd.(www.searchspace.com), which has developed a system for the London Stock Exchange called MonITARS (Monitoring Insider Trading and Regulatory Surveillance) that combines genetic algorithms, fuzzy logic, and neural network technology to detect insider dealing and market manipulation. Chartier and Spillane (2000) also describe an application of neural networks to detecting money laundering.

5. Telecommunications Fraud

The telecommunications industry has expanded dramatically in the last few years with the development of affordable mobile phone technology. With the increasing number of mobile phone users, global mobile phone fraud is also set to rise. Various estimates have been presented for the cost of this fraud. For example, Cox *et al* (1997) give a figure of $1 billion a year. *Telecom and Network Security Review* (Vol4(5), April 1997) gave a figure of between 4% and 6% of US telecoms revenue being lost due to fraud. Cahill *et al* (2002) suggests that international figures are worse, with ‘several new service providers reporting losses over 20%’. Moreau *et al* (1996) give a value of ‘several million ECUs per year’. Presumably this refers to within the EU and, given the size of the other estimates, we wonder if this should be ‘billions’. According to a recent report (Neural Technologies, 2000) ‘the industry already reports a loss of £13 billion each year due to fraud’. Mobile Europe (2000) gives a figure of US$13 billion. The latter article also claims that it is estimated that fraudsters can steal up to 5% of some operators’ revenues, and that some expect telecoms fraud as a whole to reach $28 billion per year within three years.

Despite the variety in these figures, it is clear that they are all very large. Apart from the fact that they are simply estimates, and hence subject to expected inaccuracies and variability based on the information used to derive them, there are other reasons for the differences. One is the distinction between hard and soft currency. Hard currency is real money, paid by someone other than the perpetrator for the service the perpetrator has stolen. Hynninen (2000) gives the example of the sum one mobile phone operator will pay another for the use of their network. Soft currency is the value of the service the perpetrator has stolen. At least part of this is only a loss if one assumes that the thief would have used the same service even if he or she had had to pay for it. Another reason for the differences derives from the fact that such estimates may be used for different purposes. Hynninen (2000) gives the examples of operators giving estimates on the high side, hoping for more stringent anti-fraud legislation, and operators giving estimates on the low side to encourage customer confidence.

We need to distinguish between fraud aimed *at* the service provider and fraud enabled *by* the service provider. An example of the former is the resale of stolen call time, and an example of the latter is interfering with telephone banking instructions. (It is the possibility of the latter sort of fraud which makes the public wary of using their credit cards over the Internet.) We can also distinguish between revenue fraud and non-revenue fraud. The aim of the former is to make money for the perpetrator, while the aim of the latter is simply to obtain a service free of charge (or, as with computer hackers, for example, the simple challenge represented by the system).

There are many different types of telecoms fraud (see, for example, Shawe-Taylor *et al*, 2000), and these can occur at various levels. The two most prevalent types are subscription fraud and superimposed or ‘surfing’ fraud. Subscription fraud occurs when the fraudster obtains a subscription to a service, often with false identity details, with no intention of paying. This is thus at the level of a phone number – all transactions from this number will be fraudulent. Superimposed fraud is the use of a service without having the necessary authority and is usually detected by the appearance of ‘phantom’ calls on a bill. There are several ways to carry out superimposed fraud, including mobile phone cloning and obtaining calling card authorization details. Superimposed fraud will generally occur at the level of individual calls – the fraudulent calls will be mixed in with the legitimate ones. Subscription fraud will generally be detected at some point through the billing process – though one would aim to detect it well before that, since large costs can quickly be run up. Superimposed fraud can remain undetected for a long time. The distinction between these two types of fraud follows a similar distinction in credit card fraud.

Other types of telecoms fraud include ‘ghosting’ (technology that tricks the network in order to obtain free calls) and ‘insider’ fraud where telecom company employees sell information to criminals that can be exploited for fraudulent gain. This, of course, is a universal cause of fraud, whatever the domain. ‘Tumbling’ is a type of superimposed fraud in which rolling fake serial numbers are used on cloned handsets, so that successive calls are attributed to different legitimate phones. The chance of detection by spotting unusual patterns is small, and the illicit phone will operate until all of the assumed identities have been spotted. The term ‘spoofing’ is sometimes used to describe users pretending to be someone else.

Telecommunications networks generate vast quantities of data, sometimes of the order of several gigabytes per day, so that data mining techniques are of particular importance. The 1998 database of AT&T for example, contained 350 million profiles and processed 275 million call records per day (Cortes and Pregibon, 1998).

As with other fraud domains, apart from some domain specific tools, methods for detection hinge around outlier detection and supervised classification, either using rule-based methods or based on comparing statistically derived suspicion scores with some threshold. At a low level, simple rule-based detection systems use rules such as the apparent use of the same phone in two very distant geographical locations in quick succession, calls which appear to overlap in time, and very high value and very long calls. At a higher level, statistical summaries of call distributions (often called profiles or *signatures* at the user level) are compared with thresholds determined either by experts or by application of supervised learning methods to known fraud/non-fraud cases. Murad and Pinkas (1999) and Rosset *et al* (1999) distinguish between profiling at the levels of individual calls, daily call patterns, and overall call patterns, and describe what are effectively outlier detection methods for detecting anomalous behaviour. A particularly interesting description of profiling methods is given by Cortes and Pregibon (1998). Cortes *et al* (2000) describe the *Hancock* language for writing programs for processing profiles, basing the signatures on such quantities as average call duration, longest call duration, number of calls to particular regions in the last day, and so on. Profiling and classification techniques are also described by Fawcett and Provost (1997a, 1997b, 1999) and Moreau *et al* (1997). Some work (see, for example, Fawcett and Provost, 1997a) has focused on detecting changes in behaviour.

A general complication is that signatures and thresholds may need to depend on time of day, type of account, and so on, and that they will probably need to be updated over time. Cahill *et al* (2002) suggest not including the very suspicious scores in this updating process, although more work is needed in this area.

Once again, neural networks have been widely used. The main fraud detection software of the Fraud Solutions unit of Nortel Networks (Nortel, 2000) uses a combination of profiling and neural networks. Likewise, ASPeCT (Moreau *et al* (1996), Shawe-Taylor *et al* (2000)), a project of the European Commission, Vodaphone, other European telecom companies, and academics developed a combined rule-based profiling and neural network approach. Taniguchi *et al* (1998) describe neural networks, mixture models, and Bayesian networks in telecoms fraud detection, based on call records stored for billing.

Link analysis, with links updated over time, establishes the ‘communities of interest’ (Cortes *et al*, 2001) that can indicate networks of fraudsters. These methods are based on the observation that fraudsters seldom change their calling habits, but are often closely linked to other fraudsters. Using similar patterns of transactions to infer the presence of a particular fraudster is in the spirit of phenomenal data mining (McCarthy, 2000).

Visualisation methods (Cox *et al*, 1997), developed for mining very large data sets, have also been developed for use in telecoms fraud detection. Here human pattern recognition skills interact with graphical computer display of quantities of calls between different subscribers in various geographical locations. A possible future scenario would be to code into software the patterns which humans detect.

The telecoms market will become even more complicated over time – with more opportunity for fraud. At present the extent of fraud is measured by taking account of factors such as call lengths and tariffs. The third generation of mobile phone technology will also need to take account of such things as the content of the calls (because of the packet switching technology used, equally long data transmissions may contain very different numbers of data packets), and the priority of the call.

6. Computer Intrusion

On Thursday, September 21, 2000, a sixteen-year-old boy was jailed for hacking into both the Pentagon and NASA computer systems. Between the 14th and 25th of October 2000 Microsoft security tracked the illegal activity of a hacker on the Microsoft Corporate Network. These examples illustrate that even exceptionally well protected domains can have their computer security compromised.

Computer intrusion fraud is big business and computer intrusion detection is a hugely intensive area of research. Hackers can find passwords, read and change files, alter source code, read emails, and so on. Denning (1997) lists eight kinds of computer intrusion. If the hackers can be prevented from penetrating the computer system, or can be detected early enough, then such crime can be virtually eliminated. However, as with all fraud when the prizes are high, the attacks are adaptive and once one kind of intrusion has been recognised the hacker will try a different route. Because of its importance, a great deal of effort has been put into developing intrusion detection methods, and there are several commercial products available, including CSIDS (CSIDS, 1999) and NIDES (Anderson *et al*, 1995).

Since the only record of a hacker's activities is the sequence of commands that is used when compromising the system, analysts of computer intrusion data predominantly use sequence analysis techniques. As with other fraud situations, both supervised and unsupervised methods are used. In the context of intrusion detection, supervised methods are sometimes called *misuse detection*, while the unsupervised methods used are generally methods of anomaly detection, based on profiles of usage patterns for each legitimate user. Supervised methods have the problem described in other contexts, that they can, of course, only work on intrusion patterns which have already occurred (or partial matches to these). Lee and Stolfo (1998) apply classification techniques to data from a user or program that has been identified as either 'normal' or 'abnormal'. Lippmann *et al* (2000) concluded that emphasis should be placed on developing methods for detecting new patterns of intrusion rather than old patterns, but Kumar and Spafford (1994) remark that 'a majority of break-ins…are the result of a small number of known attacks, as evidenced by reports from response teams (e.g. CERT). Automating detection of these attacks should therefore result in the detection of a significant number of break-in attempts.' Shieh and Gligor (1991, 1997) describe a pattern-matching method, and argued that it is more effective than statistical methods at detecting known types of intrusion, but is unable to detect novel kinds of intrusion patterns, which could be detected by statistical methods.

Since intrusion represents behaviour, and the aim is to distinguish between intrusion behaviour and usual behaviour in sequences, Markov models have naturally been applied (e.g. Ju and Vardi, 2001). Qu *et al* (1998) also use probabilities of events to define the profile. Forrest *et al* (1996) describe a method based on how natural immune systems distinguish between 'self' and alien patterns. As with telecoms data, both individual user patterns and overall network behaviour change over time, so that a detection system must be able to adapt to changes, but not adapt so rapidly that it also accepts intrusions as legitimate changes. Lane and Brodley (1998) and Kosoresow and Hofmeyr (1997) also use similarity of sequences that can be interpreted in a probabilistic framework

Inevitably, neural networks have been used: Ryan *et al* (1997) perform profiling by training a neural network on the process data, and also reference other neural approaches. In one of the more careful studies in the area, Schonlau *et al* (2001) describe a comparative study of six statistical approaches for detecting impersonation of other users ('masquerading'), where they took real usage data from 50 users and planted contaminating data from other users to serve as the masquerade targets to be detected. A nice overview of statistical issues in computer intrusion detection is given by Marchette (2001), and the October 2000 edition of Computer Networks (Volume 34, Issue 4) is a special issue on (relatively) recent advances in intrusion detection systems , including several examples of new approaches to computer intrusion detection.

7. Medical and Scientific Fraud

Medical fraud can occur at various levels. It can occur in clinical trials (see, for example, Buyse *et al* (1999)). It can also occur in a more commercial context – for example, prescription fraud, claiming for patients that are dead or who do not exist, and 'upcoding' where a doctor performs a medical procedure but charges the insurer for one that is more expensive, or perhaps does not even perform one at all. Allen (2000) gives an example of bills being submitted for more than 24 hours in a working day. He, Wang, *et al* (1997) and He, Graco, *et* al (1997) describe the use of neural networks, genetic algorithms, and nearest neighbour methods to classify the practice profiles of general practitioners in Australia into classes from normal to abnormal.

Medical fraud is often linked to insurance fraud: Terry Allen, a statistician with the Utah Bureau of Medicaid Fraud, reports an estimate of up to 10% of the $800 million annual claims may be stolen (Allen 2000). Major and Riedinger (1992) created a 'knowledge/statistical-based system' for detecting healthcare fraud by comparing observations with those with which they should be most similar (for example, having similar geodemographics). Brockett *et al* (1998) use neural networks to classify fraudulent and non-fraudulent claims for automobile bodily injury in healthcare insurance claims. Glasgow (1997) gives a short discussion of risk and fraud in the insurance industry. A glossary of several of the different types of medical fraud is given at http://www.motherjones.com/mother_jones/MA95/davis2.html.

Of course, medicine is not the only scientific area where data have sometimes been fabricated, falsified, or carefully selected to support a pet theory. Problems of fraud in science are attracting increased attention, but they have always been with us: errant scientists have been known to massage figures from experiments in order to push through development of a product or reach a magical 'significance level' for a publication. Dmitriy Yuryev describes such a case on his WebPages at http://www.orc.ru/~yur77/statfr.htm. Moreover, there are many classical cases in which the data have been suspected of being massaged (including the work of Galileo, Newton, Babbage, Kepler, Mendel, Millikan, and Burt). Press and Tanur (2001) present a fascinating discussion of the role of subjectivity in the scientific process, illustrating with many examples. The borderline between subconscious selection of data and out-and-out distortion is a fine one.

8. Conclusions

The areas we have outlined above are perhaps those in which statistical and other data analytic tools have made most impact on fraud detection. This is typically because there are large quantities of information, and this information is numerical or can easily be converted into the numerical in the form of counts and proportions. However, other areas, not mentioned above, have also used statistical tools for fraud detection. Irregularities in financial statements can be used to detect accounting and management fraud in contexts broader than those of money laundering. Digit analysis tools have found favour in accountancy (e.g. Nigrini and Mittermaier, 1997; Nigrini 1999). Statistical sampling methods are important in financial audit, and screening tools are applied to decide which tax returns merit detailed investigation. We mentioned insurance fraud in the context of medicine above, but it clearly occurs more widely. Artís *et al* (1999) describe an approach to modelling fraud behaviour in car insurance, and Fanning *et al* (1995) and Green and Choi (1997) examine neural network classification methods for detecting management fraud. Statistical tools for fraud detection have also been applied to sporting events. For example, Robinson and Tawn (1995), Smith (1997), and Barao and Tawn (1999) examined the results of running events to see if some exceptional times were out of line with what might be expected.

Plagiarism is also a type of fraud. We briefly referred to the use of statistical tools for author verification, and such methods can be applied here. However, statistical tools can also be applied more widely. For example, with the evolution of the Internet it is extremely easy for students to plagiarise articles and pass them off as their own in school or university coursework. The website http://www.plagiarism.org describes a system that can take a manuscript and compare it against their ‘substantial database’ of articles from the Web. A statistical measure of the originality of the manuscript is returned.

As we commented in the introduction, fraud detection is a post hoc strategy, being applied after fraud prevention has failed. Statistical tools are also applied in some fraud prevention methods. For example, so-called *biometric* methods of fraud detection are slowly becoming more widespread. These include computerised fingerprint and retinal identification, and also face recognition (though this has received most publicity in the context of recognising football hooligans).

In many of the applications we have discussed, speed of processing is of the essence. This is particularly the case in transaction processing, especially with telecoms and intrusion data, where vast numbers of records are processed every day, but also applies in credit card, banking, and retail sectors.

A key issue in all of this work is how effective the statistical tools are in detecting fraud and a fundamental problem is that one typically does not know how many fraudulent cases slip through the net. In applications such as banking fraud and telecoms fraud, where speed of detection matters, measures such as average time to detection after fraud starts (in minutes, numbers of transactions, etc.) should also be reported. Measures of this aspect interact with measures of final detection rate: in many situations an account, telephone, etc., will have to be used for several fraudulent transactions before it is detected as fraudulent, so that several false negative classifications will necessarily be made.

An appropriate overall strategy is to use a graded system of investigation. Accounts with very high suspicion scores merit immediate and intensive (and expensive) investigation, while those with large but less dramatic scores merit closer (but not expensive) observation. Once again, it is a matter of choosing a suitable compromise.

Finally, it is worth repeating the conclusions reached by Schonlau *et al* (2001), in the context of statistical tools for computer intrusion detection: that ‘statistical methods can detect intrusions, even in difficult circumstances’, but also that ‘many challenges and opportunities for statistics and statisticians remain’. We believe this positive conclusion holds more generally. Fraud detection is an important area, one in many ways ideal for the application of statistical and data analytic tools, and one where statisticians can make a very substantial and important contribution.

Acknowledgements

The work of Richard Bolton was supported by a ROPA award from the Engineering and Physical Sciences Research Council of the UK.

References

Aleskerov, E., Freisleben, B., and Rao, B. (1997). CARDWATCH: a neural network based database mining system for credit card fraud detection. *Computational Intelligence for Financial Engineering, Proceedings of the IEEE/IAFE*, 220-226.

Allen, T. (2000), A day in the life of a Medicaid fraud statistician. *STATS* **29**, 20-22.

Anderson, D., Frivold, T., and Valdes, A. (1995). Next-generation intrusion detection expert system (NIDES): a summary. *Technical Report SRI-CSL-95-07*, Computer Science Laboratory, SRI International.

Andrews, P.P. and Peterson M.B. (eds.) (1990). *Criminal Intelligence Analysis*. Loomis, California: Palmer Enterprises.

Artís, M., Ayuso, M. and Guillén, M. (1999). Modelling different types of automobile insurance fraud behaviour in the Spanish market. *Insurance Mathematics and Economics* **24**, 67-81.

Barao, M. I. and Tawn, J. A. (1999). Extremal analysis of short series with outliers: sea-levels and athletics records. *Journal of the Royal Statistical Society Series C-Applied Statistics* **48**, 469-487.

Blunt, G. and Hand, D. J. (2000). *The UK credit card market*. Technical Report, Department of Mathematics, Imperial College, London.

Bolton, R. J. and Hand, D. J. (2001). Unsupervised profiling methods for fraud detection, *Credit Scoring and Credit Control VII, Edinburgh, UK, 5-7 Sept.*

Brause, R., Langsdorf, T. and Hepp, M. (1999). Neural data mining for credit card fraud detection. *Proceedings. 11th IEEE International Conference on Tools with Artificial Intelligence*.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and regression trees*. Belmont, California, U.S.A., Wadsworth Publishing Company.

Brockett, P. L., Xia, X. and Derrig, R. A. (1998). Using Kohonen's self-organising feature map to uncover automobile bodily injury claims fraud. *The Journal of Risk and Insurance* **65**(2), 245-274.

Burge, P. and Shawe-Taylor, J. (1997). Detecting cellular fraud using adaptive prototypes. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, 9-13.

Buyse, M., George, S. L., Evans, S., Geller, N. L., Ranstam, J., Scherrer, B., Lesaffre, E., Murray, G., Edler, L., Hutton, J., Colton, T., Lachenbruch, P. and Verma, B. L. (1999). The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Statistics in Medicine* **18**, 3435-3451.

Cahill, M.H., Lambert, D., Pinheiro, J.C., and Sun, D.X. (2002). Detecting fraud in the real world. To appear in *Handbook of Massive Datasets*, J. Abello, P. M. Pardalos, M. G. C. Resende (eds.), Klewer.

Chan, P. and Stolfo, S. (1998). Toward scalable learning with non-uniform class and cost distributions: A case study in credit card fraud detection. *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*, 164-168.

Chan, P. K., Fan, W., Prodromidis, A. L. and Stolfo, S. J. (1999). Distributed data mining in credit card fraud detection. *IEEE Intelligent Systems* **14**(6), 67-74.

Chartier, B. and Spillane, T. (2000). Money laundering detection with a neural network. *Business Applications of Neural Networks*. P.J.G. Lisboa, A.Vellido, B.Edisbury Eds. Singapore: World Scientific. 159-172.

Chhikara, R. S. and McKeon, J. (1984). Linear discriminant analysis with misallocation in training samples. *JASA*, **79**, 899-906.

Clark, P., and Niblett, T. (1989). The CN2 induction algorithm. Machine Learning 3, 261-285.

Cohen, W. (1995). Fast effective rule induction. *Proceedings 12th International Conference of Machine Learning*, 115-123.

Cortes, C. and Pregibon, D. (1998). Giga-mining. *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining,* 174-178.

Cortes, C., Fisher, K., Pregibon D. and Rogers A. (2000). Hancock: a language for extracting signatures from data streams. *Proceedings of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, N.Y., 9-17.

Cortes, C, Pregibon D. and Volinsky C. (2001). Communities of interest. *Proceedings IDA2001*, Lisbon, Portugal.

Cox, K. C., Eick, S. G., Wills, G. J. and Brachman, R. J. (1997). Visual data mining: recognizing telephone calling fraud. *Journal of Data Mining and Knowledge Discovery* **1**(2), 225-231.

CSIDS (1999). Cisco secure intrusion detection system technical overview. http://www.wheelgroup.com/warp/public/cc/cisco/mkt/security/nranger/tech/ntran_tc.htm

Denning, D.E. (1997). Cyberspace attacks and countermeasures. In *Internet Besieged*. Ed. D.E.Denning and P.J.Denning, New York: ACM Press. 29-55.

Dorronsoro, J. R., Ginel, F., Sanchez, C. and Cruz, C. S. (1997). Neural fraud detection in credit card operations. *IEEE Transactions on Neural Networks* **8**(4), 827-834.

Fanning, K., Cogger, K. O. and Srivastava, R. (1995). Detection of management fraud: a neural network approach. *Journal of Intelligent Systems in Accounting, Finance and Management* **4**, 113-126.

Fawcett, T. and Provost, F. (1997a). Adaptive fraud detection. *Data Mining and Knowledge Discovery* **1**(3), 1-28.

Fawcett, T. and Provost, F. (1997b). Combining data mining and machine learning for effective fraud detection. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press, Menlo Park, California, 14-19.

Fawcett, T. and Provost, F. (1999). Activity monitoring: noticing interesting changes in behavior. *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, N.Y, 53-62.

Forrest, S., Hofmeyr, S., Somayaji, A. and Longstaff, T. (1996). A sense of self for Unix processes. *Proceedings of the 1996 IEEE Symposium on Security and Privacy*, Los Alamitos, CA., 120-128.

Ghosh, S. and Reilly, D. L. (1994). Credit card fraud detection with a neural-network. *Proceedings of the 27th Annual Hawaii International Conference on System Science. Volume 3 : Information Systems: DSS/Knowledge-Based*

Systems, J. F. Nunamaker and R. H. Sprague, Eds., Los Alamitos, CA, USA, 621-630.

Glasgow, B. (1997). Risk and fraud in the insurance industry. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management,* AAAI Press, Menlo Park, California, 20-21.

Goldberg, H. and Senator, T. E. (1995). Restructuring databases for knowledge discovery by consolidation and link formation. *Proceedings of the 1st Internationall Conference on Knowledge Discovery and Data Mining,* 136-141.

Goldberg, H. and Senator, T. E. (1997). Break detection systems. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press, Menlo Park, California, 22-28.

Goldberg, H. and Senator, T. E. (1998). The FinCEN AI system: finding financial crimes in a large database of cash transactions. *Agent Technology: Foundations, Applications, and Markets*. ed N.Jennings and M.Wooldridge, Springer-Verlag, Berlin, 283-302.

Green, B. P. and Choi, J. H. (1997). Assessing the risk of management fraud through neural network technology. *Auditing* **16**(1), 14-28.

Hand, D.J. (1981). *Discrimination and Classification*. Chichester: Wiley.

Hand, D.J. (1997). *Construction and Assessment of Classification Rules*. Chichester: Wiley.

Hand, D.J. and Blunt, G. (2001). Prospecting for gems in credit card data. *IMA Journal of Management Mathematics*, **12**, 173-200.

Hand, D.J., Blunt, G., Kelly, M.G., and Adams, N.M. (2000). Data mining for fun and profit. *Statistical Science* **15**, 111-131.

Hand, D.J. and Henley, W.E. (1997). Statistical classification methods in consumer credit scoring: a review. *Journal of the Royal Statistical Society, Series A* **160**, 523-541.

Hassibi, K. (2000). Detecting payment card fraud with neural networks. *Business Applications of Neural Networks*. P.J.G. Lisboa, A.Vellido, B.Edisbury Eds. Singapore: World Scientific.

He, H., Graco, W. and Yao, X. (1999). Application of genetic algorithm and *k*-nearest neighbour method in medical fraud detection. *Lecture Notes in Computer Science* **1585**, 74-81.

He, H. X., Wang, J. C., Graco, W. and Hawkins, S. (1997). Application of neural networks to detection of medical fraud. *Expert Systems with Applications* **13**(4), 329-336.

Hill, T. P. (1996). A statistical derivation of the significant-digit law. *Statistical Science* **10**(4), 354-363.

Hynninen, J. (2000). Experiences in mobile phone fraud. Tik-110.501 Seminar on Network Security, TML, Helsinki University of Technology, Finland.

Jenkins P. (2000). Getting smart with fraudsters. *Financial Times*, September 23.

Jensen, D. (1997). Prospective assessment of AI technologies for fraud detection: a case study. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press, Menlo Park, California, 34-38.

Ju, W-H and Vardi, Y. (2001). A hybrid high-order Markov chain model for computer intrusion detection. *Journal of Computational and Graphical Statistics* **10**, 277-295.

Kirkland, J. D., Senator, T. E., Hayden, J. J., Dybala, T., Goldberg, H. G. and Shyr, P. (1998). The NASD regulation Advanced Detection System (ADS). *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-98) and of the 10th Conference on Innovative Applications of Artificial Intelligence (IAAI-98)*, Menlo Park, 1055-1062.

Kosoresow, A. P. and Hofmeyr, S. A. (1997). Intrusion detection via system call traces. *IEEE Software* **14**(5), 24-42.

Kumar, S. and Spafford, E. (1994). A pattern matching model for misuse intrusion detection. *Proceedings of the 17th National Computer Security Conference*, 11-21.

Lachenbruch, P. A. (1966) Discriminant analysis when the initial samples are misclassified. *Technometrics*, **8**, 657-662.

Lachenbruch, P. A. (1974) Discriminant analysis when the initial samples are misclassified II: non-random misclassification models. *Technometrics*, **16**, 419-424.

Lane, T. and Brodley, C. E. (1998). Temporal sequence learning and data reduction for anomaly detection. *Proceedings of the 5th ACM Conference on Computer and Communications Security (CCS-98)*. New York, 150-158.

Lee, W. and Stolfo, S. (1998). Data mining approaches for intrusion detection. *Proceedings of the 7th USENIX Security Symposium*. San Antonio, TX, 79-93.

Leonard, K. J. (1993). Detecting credit card fraud using expert systems. *Computers and Industrial Engineering* **25**(1-4), 103-6.

Lippmann, R., Fried, D., Graf, I., Haines, J., Kendall, K., McClung, D., Weber, D., Webster, S., Wyschogrod, D., Cunningham, R., and Zissman, M. (2000). Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion-detection evaluation. MIT Lincoln Laboratory. Unpublished manuscript.

Major, J. A. and Riedinger, D. R. (1992). EFD: A hybrid knowledge/statistical-based system for the detection of fraud. *International Journal of Intelligent Systems* **7**, 687-703.

Marchette, D.J. (2001) *Computer Intrusion Detection and Network Monitoring: a Statistical Viewpoint*. New York: Springer.

McCarthy, J. (2000). Phenomenal Data Mining. *SIGKDD Explorations* **1**(2), 24-29.

McLachlan G.J. (1992) *Discriminant Analysis and Statistical Pattern Recognition*. New York: John Wiley and Sons.

Mobile Europe (2000). New IP World, new dangers. *Mobile Europe*, March.

Moreau, Y., Verrelst, H. and Vandewalle, J. (1997). Detection of mobile phone fraud using supervised neural networks: a first prototype. In *International Conference on Artificial Neural Networks Proceedings (ICANN'97)*, 1065-1070.

Moreau, Y., B. Preneel, P. Burge, J. Shawe-Taylor, C. Stoermann and C. Cooke (1996). Novel techniques for fraud detection in mobile telecommunications. *ACTS Mobile Summit, Grenada*.

Murad, U., and Pinkas, G. (1999). Unsupervised profiling for identifying superimposed fraud. In *Principles of Data Mining and Knowledge Discovery, Lecture Notes in Artificial Intelligence* **1704**, 251-261.

Neural Technologies (2000). *Reducing telecoms fraud and churn*. Neural Technologies Ltd, Bedford Road, Petersfield, Hampshire GU32 3QA.

Nigrini, M.J. and Mittermaier, L.J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing: A Journal of Practice and Theory* **16**, 52-67.

Nigrini, M.J. (1999). I've got your number. *Journal of Accountancy*, May, 79-83.

Nortel (2000) Nortel Networks Fraud Solutions (2000). Fraud Primer. Issue 2.0. Nortel Networks Corporation.

Pak, S.J. and Zdanowicz, J.S. (1994). A statistical analysis of the U.S. Merchandise Trade Database and its uses in transfer pricing compliance and enforcement. *Tax Management*, (US Bureau of National Affairs, Inc.), May 11.

Patient, S. (2000). Reducing Online Credit Card Fraud, *Web Developer's Journal*, http://www.webdevelopersjournal.com/articles/card_fraud.html

Press, S.J. and Tanur, J.M. (2001). *The Subjectivity of Scientists and the Bayesian Approach*. New York: Wiley.

Provost, F. and Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning* **42**(3), 203-231.

Qu, D., Vetter, B. M., Wang, F., Narayan, R., Wu, S. F., Hou, Y. F., Gong, F. and Sargor, C. (1998). Statistical Anomaly Detection for Link-State Routing Protocols. *Proceedings. Sixth International Conference on Network Protocols*, 62-70.

Quinlan, J. R. (1990). Learning logical definitions from relations. *Machine Learning* **5**, 239-266.

Quinlan, J. R. (1993). *C4.5: programs for machine learning*. San Mateo, CA, Morgan Kaufmann.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge University Press.

Robinson, M. E. and Tawn, J. A. (1995). Statistics for exceptional athletics records. *Applied Statistics* **44**(4), 499-511.

Rosset, S., Murad, U., Neumann, E., Idan, Y., and Pinkas, G. (1999). Discovery of fraud rules for telecommunications – challenges and solutions. In

Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 409-413.

Ryan, J., Lin, M. and Miikkulainen, R. (1997). Intrusion detection with neural networks. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press, Menlo Park, California, 72-79.

Schonlau, M., DuMouchel, W., Ju, W. H., Karr, A. F., Theus, M., Vardi, Y. (2001). Computer intrusion: detecting masquerades. *Statistical Science* **16**(1), 58-74.

Senator, T. E., Goldberg, H. G., Wooton, J., Cottini, M. A., Khan, A. F. U., Klinger, C. D., Llamas, W. M., Marrone, M. P. and Wong, R. W. H. (1995). The Financial Crimes Enforcement Network AI system (FAIS) – Identifying potential money laundering from reports of large cash transactions. *AI Magazine* **16**(4), 21-39.

Senator, T. E. (2000). Ongoing management and application of discovered knowledge in a large regulatory organization: a case study of the use and impact of NASD regulation's Advanced Detection System (ADS). *Proceedings of the sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. N.Y., 44-53.

Shawe-Taylor, J., Howker, K., Gosset, P., Hyland, M., Verrelst, H., Moreau, Y., Stoermann, C. and Burge, P. (2000). Novel techniques for profiling and fraud detection in mobile telecommunications. In *Business Applications of Neural*

Networks. P.J.G. Lisboa, A.Vellido, B.Edisbury Eds. Singapore: World Scientific, 113-139.

Shieh, S.-P. W. and Gligor, V. D. (1991). A pattern oriented intrusion model and its applications. *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, 327-342.

Shieh, S.-P. W. and Gligor, V. D. (1997). On a pattern-oriented model for intrusion detection. *IEEE Transactions on Knowledge and Data Engineering* **9**(4), 661-667.

Smith, R.L. (1997). Statistics for exceptional athletics records. *Applied Statistics* **46** 123-128.

Stolfo, S. J., Fan, D. W., Lee, W., Prodromidis, A. L. and Chan, P. K. (1997). Credit card fraud detection using meta-learning: issues and initial results. *AAAI Workshop: AI Approaches to Fraud Detection and Risk Management*, AAAI Press, Menlo Park, California, 83-90.

Stolfo, S., Fan, W., Lee, W., Prodromidis, A. L., and Chan, P. (2000). Cost-based modeling for fraud and intrusion detection: results from the JAM Project. *DARPA Information Survivability Conference and Exposition*, IEEE Computer Press.

Stolfo, S. J., Prodromidis, A. L., Tselepis, S., Lee, W., Fan, D. W. and Chan, P. K. (1997). JAM: Java Agents for Meta-learning over distributed databases. *AAAI*

Workshop: AI Approaches to Fraud Detection and Risk Management, AAAI Press, Menlo Park, California, 91-98.

Taniguchi, M., Haft, M., Hollmén, J, and Tresp, V. (1998). Fraud detection in communication networks using neural and probabilistic methods. In *Proceedings of the 1998 IEEE International Conference in Acoustics, Speech and Signal Processing (ICASSP'98), Volume 2*, 1241-1244.

U.S. Congress (1995). *Information technologies for the control of money laundering*. OTA-ITC-630, US Congress, Office of Technology Assessment, Washington, DC: US Government Printing Office.

Wasserman, S., and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Webb, A.R. (1999). *Statistical Pattern Recognition*, London: Arnold.

Wheeler, R. and Aitken, S. (2000). Multiple algorithms for fraud detection. *Knowledge-Based Systems* **13**(2-3), 93-99.