Accepted Manuscript

Title: Intelligent financial fraud detection: a comprehensive review

Author: Jarrod West, Maumita Bhattacharya

PII: S0167-4048(15)00126-1
DOI: http://dx.doi.org/doi:10.1016/j.cose.2015.09.005
Reference: COSE 941

To appear in: *Computers & Security*

Received date: 11-9-2014
Revised date: 10-4-2015
Accepted date: 8-9-2015

Please cite this article as: Jarrod West, Maumita Bhattacharya, Intelligent financial fraud detection: a comprehensive review, *Computers & Security* (2015), http://dx.doi.org/doi:10.1016/j.cose.2015.09.005.

Title of the paper: Intelligent Financial Fraud Detection: A Comprehensive Review

Authors:

Jarrod West
School of Computing & Mathematics
Charles Sturt University
Australia – 2640

Maumita Bhattacharya
School of Computing & Mathematics
Charles Sturt University
Australia – 2640

Corresponding author:

Maumita Bhattacharya

Phone: +61 2 60519619
Fax: +61 2 60519619
Email: maumita.bhattacharya@ieee.org

Postal address:

Maumita Bhattacharya
Honours Course Coordinator
School of Computing & Mathematics
Charles Sturt University, Albury Campus
PO Box: 789
Albury, NSW
Australia - 2640

Author Biography

Jarrod West

Jarrod West is currently studying Honous in the School of Computing & Mathematics, Charles Sturt University, Australia. Jarrod completed his Bachelor degree in Computer Science from the University of Newcastle, Australia in 2013. As a researcher he is primarily interested in various aspects of information security. Jarrod's current research is focused on computational intelligence and data mining based financial fraud detection.

Maumita Bhattacharya

Maumita Bhattacharya currently works as an academic in the School of Computing & Mathematics, Charles Sturt University, Australia. As an academic Maumita has been involved in research and teaching in various disciplines of computer science and engineering including artificial intelligence, data mining, information security and computer ethics. Her current research interests focus on design and development of advanced computational intelligence algorithms and methods and their applications to information security, data mining, optimization and other complex real world problem domains. Maumita serves in the editorial boards and scientific committees of many reputed journal and conferences in these fields.

Intelligent Financial Fraud Detection: A Comprehensive Review

Jarrod West and Maumita Bhattacharya

School of Computing & Mathematics, Charles Sturt University
NSW-2640, Australia

Abstract. Financial fraud is an issue with far reaching consequences in the finance industry, government, corporate sectors, and for ordinary consumers. Increasing dependence on new technologies such as cloud and mobile computing in recent years has compounded the problem. Traditional methods involving manual detection are not only time consuming, expensive and inaccurate, but in the age of big data they are also impractical. Not surprisingly, financial institutions have turned to automated processes using statistical and computational methods. This paper presents a comprehensive review of financial fraud detection research using such data mining methods, with a particular focus on computational intelligence (CI)-based techniques. Over fifty scientific literature, primarily spanning the period 2004-2014, were analysed in this study; literature that reported empirical studies focusing specifically on CI-based financial fraud detection were considered in particular. Research gap was identified as none of the existing review articles addresses the association among fraud types, CI-based detection algorithms and their performance, as reported in the literature. We have presented a comprehensive classification as well as analysis of existing fraud detection literature based on key aspects such as detection algorithm used, fraud type investigated, and performance of the detection methods for specific financial fraud types. Some of the key issues and challenges associated with the current practices and potential future direction of research have also been identified.

Key words: Financial fraud detection; Computational intelligence; Data mining; Anomaly detection; Classification

1 Introduction

Financial fraud is an issue that has wide reaching consequences in both the finance industry and daily life. Fraud can reduce confidence in industry, destabilise economies, and affect people's cost of living. Traditional approaches relied on manual techniques such as auditing, which are inefficient and unreliable due to the difficulty of the problem. Data mining-based approaches have been shown to be useful because of their ability to identify small anomalies in large data sets [32]. There are many

different types of fraud, as well as a variety of data mining methods, and research is continually being undertaken to find the best approach for each case.

Financial fraud is a broad term with various potential meanings, but for our purposes it can be defined as the intentional use of illegal methods or practices for the purpose of obtaining financial gain [59]. Fraud has a large negative impact on business and society: credit card fraud alone accounts for billions of dollars of lost revenue each year [7], and some figures suggest that the total yearly cost to the U.S. could be in excess of $400 billion [28], while a third study shows that UK insurers are out 1.6 billion pounds a year due to fraudulent claims [32]. Financial fraud also has broader ramifications on the industry, such as providing funding for illicit activities like drug trafficking and organised crime [7]. For credit card fraud the cost is typically worn by the merchants, who end up paying shipping, chargeback, and administrative costs as well as losing consumer confidence after being victim to a fraudulent transaction [37], [41]. In this way we can see the widespread consequences that fraud can have and the importance in minimising it.

Advancements in modern technologies such as the internet and mobile computing have led to an increase in financial fraud in recent years [54]. Social factors such as the increased distribution of credit cards has increased spending but also resulted in an increase to fraud [41]. Fraudsters are continually refining their methods, and as such there is a requirement for detection methods to be able to evolve accordingly [7]. Data mining has already been shown to be useful in similar domains such as credit card approval, bankruptcy prediction, and analysis of share markets [35]. Fraud detection is a considered to be a similar classification problem but with a vast imbalance in fraudulent to legitimate transactions, and a sizable difference in cost for misclassifying them [16]. Data mining approaches are also applicable to fraud detection for their efficiency at processing large datasets and their ability to work without requiring knowledge of the input variables [38].

A useful framework for applying data mining to fraud detection is to use it as a method for classifying suspicious transactions or samples for further consideration. Studies show that reviewing 2% of credit card transactions could reduce fraud losses to 1% of the total cost of all purchases, with more assessments resulting in smaller loss but with an increase in auditing costs [37]. A multi-layer pipeline approach can be used with each step applying a more rigorous method to detect fraud. Data mining can be utilised to efficiently filter out more obvious fraud cases in the initial levels and leave the more subtle ones to be reviewed manually [37].

In this article we will use a few broad terminologies that are defined here for clarity. Data mining refers to any method that processes large quantities of data to derive an underlying meaning. Within this classification we will consider two categories of data mining: statistical and computational. We define the statistical techniques as those that are based on traditional mathematical methods, such as logistic regression and Bayesian theory. Computational methods are those which use modern intelligence techniques, such as neural networks and support vector machines. Though these categories share many similarities, we will consider that the main difference between them is that computational methods are capable of learning from and adapt-

ing to the problem domain, while statistical methods are more rigid. Both types of data mining will be researched in this article.

The objective of this paper is to provide a review of existing literature in financial fraud detection and compare their findings (focusing primarily on literature published during 2004-2014 and research that reports empirical studies using CI-based techniques). A key focus of this review is on the reported performance of CI techniques for specific fraud types. None of the existing reviews (for example, [32], [62] and [63]) covers this aspect. This will provide an indication to future researchers on the areas that are currently available for future study. The remainder of the article is structured as follows. In the next section we will detail the history of intelligent fraud detection research. Section 3 will list the different types of financial fraud. Section 4 presents an overview of the assorted detection methods used. Section 5 details the specific efforts of previous researchers in detecting financial fraud. Section 6 offers an insight into what is missing from existing techniques and proposes areas for future research. Section 7 provides a conclusion to our research and a discussion of our findings.

2 Related Work

Prior investigation has already been performed on some aspects of intelligent financial fraud detection, and a brief history of the specific research methods undertaken is given here. Figure 1 shows the timeline of research for the last decade.

Initial fraud detection studies focussed heavily on statistical models such as logistic regression, as well as neural networks [37], [55], [24]. Zhang et al. discovered that neural networks had been used for financial applications such as forecasting since 1988 [58]. In 1995 Sohl and Venkatachalam first predicted financial statement fraud using a back-propagation neural network [42]. Fraser et al. compared techniques across a quantitative spectrum including statistical and computational methods such as regression and neural networks [19]. In 1998 Fanning and Cogger used a neural network based on other financial ratios and variables and found it compared favourably to discriminant analysis and logistic regression [17]. In 2001 and 2002 Bolton and Hand performed some general analysis on fraud detection, focussing specifically on statistical learning [8], [9], and Rezaee investigated financial statement fraud in depth [39]. Table 1 shows the bias towards neural networks and statistical methods with a list of further fraud detection research performed outside of the last decade.

Recent fraud detection research has been far more varied in methods studied, although the former techniques are still popular. In 2004 Kou et al. reviewed the study

of general fraud detection using analytic techniques including neural networks [30]. Vatsa et al. investigated a novel approach using game theory in 2005, which modelled fraudsters and detection methods as opposing players in a game, each striving to obtain the greatest financial advantage [48]. The following year Yang and Hwang studied health care fraud using a process-mining approach [53].

In 2007 Pinquet et al. and Viaene et al. both studied logistic regression with insurance fraud, concentrating on a database of Spanish automobile insurance claims [36], [49]. Kirkos et al. compared statistical methods with neural networks to identify fraudulent Greek manufacturing companies [28], and Bose and Wang focussed on classification and regression trees to solve financial statement fraud in a selection of Chinese companies [10].

Also in 2007 Hoogs et al. introduced a genetic algorithm on Accounting and Auditing Enforcement Releases to detect fraudulent companies in the US [24], and Yue et al. performed a review of existing fraud detection literature. They claimed that the only successful methods of fraud detection to date, as well as the most commonly researched, were classification based [55].

Bai et al. used decision trees to study financial statement fraud for a selection of Chinese companies in 2008 [3]. Bermúdez et al. took a statistical approach to insurance fraud detection, using the same samples that Pinquet et al. and Viaene et al. used previously [6]. Quah and Sriganesh looked at visualising credit card fraud with self-organising maps, focussing on real-world samples from the Singaporean branch of an international bank [37]. Wu and Banzhaf modified the standard artificial immune system method with a coevolutionary approach, using it to solve transactional fraud with automatic teller and point-of-sale data for a financial institution [52].

In 2009 Holton utilised a combination of text mining and Bayesian belief networks to identify disgruntled employees likely to commit corporate fraud [23]. Panigrahi et al. combined a Dempster-Schaefer adder with a Bayesian learner to solve credit card fraud with their own synthesised data [35]. Sánchez et al. focussed on credit cards provided by a multinational department store, using self-organising maps to cluster and visualise fraudulent patterns [41]. Whitrow et al. compared support vector machines with decision trees in solving credit card fraud, with a focus on aggregating common transactional variables to create new inputs [50].

In 2010 Cecchini et al. studied Accounting and Auditing Enforcement Releases (AAER) with their own text mining and support vector machine hybrid to predict financial statement fraud in US companies [13]. Then in 2011 Bhattacharyya et al. compared the ability of logistic regression, support vector machines, and random forests on a large sample of credit card transactions to identify which were fraudulent [7]. Duman and Ozcelik combined the strengths of genetic algorithms and scatter search to create their own hybrid method. They used it to track consumer spending with a large Turkish bank, as an aid to predicting the occurrence of credit card fraud [16].

Humphrys et al. created text mining hybrids by utilising other common methods to act as the classifier. With a support vector machine, decision tree, and Bayesian belief network they managed to successfully identify fraud within the company's 10-K document filings [26]. Glancy and Yadav also studied sections of 10-K documents

for US companies known to be fraudulent, processing the text with a singular validation decomposition vector to classify the samples [20]. Jans et al. applied process mining to the internal logs created by a European financial institution to detect corporate fraud [27], and Ngai et al. performed a substantial review of existing fraud detection literature as of 2011 [32].

Also in 2011 Ravisankar et al. compared a large range of methods to identify financial statement fraud within Chinese companies. In addition to support vector machines they looked at genetic programming, logistic regression, group method of data handling, and variety of neural networks [38]. Zhou and Kapoor created a generic framework for financial statement fraud detection using response surface methodology[59], then in 2012 Wong et al. utilised an artificial immune system to predict credit card fraud for a major Australian bank [51].

In 2013 Huang investigated financial statement fraud in a series of Taiwanese companies using logistic regression and a support vector machine [25]. Zaki and Theodoulidis took a more direct approach and focussed on the litigation section of the Securities and Exchange Commission website, applying their own text mining algorithm to classify financial statement fraud [56]. Sahin et al. studied the ability of decision trees to identify fraudulent credit card transactions, using a six-month sample from a major bank [40].

In 2014 Dong et al. used text mining to study AAERs for Chinese companies that were trading publicly in the US [15]. Olszewski visualised credit card fraud with self organising maps, focusing only on accounts held by residents of Warsaw, Poland [33]; Soltani Halvaiee and Akbari utilised an artificial immune system to identify credit card fraud for an anonymous Brazilian bank [43]; and West et al. investigated the present state of fraud detection research [60].

3 Types of Financial Fraud

There are many different types of financial fraud, and a brief description of some of the major types will be listed here (also see Figure 2). The fraud types were selected from the list provided in the Federal Bureau of Investigation, Financial Crimes Report (2010-2011), United States [61].

Credit Card Fraud. Credit card fraud refers to the unauthorised use of a person's credit card to perform fraudulent transactions without the user's knowledge [32]. The transactions can be performed using the physical card, where the card was either lost or stolen, but is often performed remotely [7]. The cardholder's information may be acquired by one of a few methods. Phishing involves a fraudster impersonating a finance official to convince the user to divulge their details, swipers or skimmers provide an interface to an ATM or POS device which can read the card directly, or entire databases of user's information can be obtained if the fraudster is able to breach the financial institutions network security or enlist the help of an accomplice within

the company [7], [37]. Obtaining the user's card could even be as simple as intercepting mail containing a new or replacement card [37]. The anonymity and availability of these remote methods has given rise to the prevalence of organised crime in credit card fraud [7]. A typical method for identifying credit card fraud is to analyse a customer's regular spending habits and flag transactions which are noticeably outside of this model [16].

Securities and Commodities Fraud. Securities fraud, also known as commodities fraud, refers to a variety of methods by which a person is deceived into investing into a company based on false information. It includes Pyramid Schemes, Ponzi Schemes, Hedge Fund Fraud, Foreign Exchange Fraud, and Embezzlement [32].

Financial Statement Fraud. Financial statements are the documents release by a company that explain details such as their expenses, loans, income and profits [38]. They can also include comments from management on the business' performance and expected issues that may arise in the future [20]. The various financial statements that the company releases give an overall picture of the company's status, and can be used to indicate how successful the company is, influence stock prices, and determine if they are applicable for loans [38]. Financial statement fraud, also known as corporate fraud, involves doctoring these statements to make the company appear more profitable.

Reasons for committing financial statement fraud include improving stock performance, reducing tax obligations, or as an attempt to exaggerate performance due to managerial pressure [38]. Financial statement fraud can be difficult to diagnose because of a general lack of understanding of the field, the infrequency in which it occurs, and the fact that it is usually committed by knowledgeable people within the industry who are capable of masking their deceit [31].

Insurance Fraud. Insurance fraud is fraud that can be committed at any point during the insurance process, and by any people in the chain. Insurance claims fraud occurs when a customer submits a fraudulent insurance claim as a result of an exaggerated injury or loss of assets, or a completely fraudulent event. A common form of claims fraud is automobile insurance fraud, which is often committed by faking or intentionally committing accidents that result in excessive repair and injury costs. Larger scale claims fraud also occurs, such as crop insurance fraud where a consumer overstates their losses due to declining agricultural prices or the effects of natural disasters. Insurance fraud can also include excessive billing, duplicate claims, kickbacks to brokers and “upcoding” of items [32].

Mortgage Fraud. Mortgage fraud is a specific form of financial fraud that refers to manipulation of a property or mortgage documents. It is often committed to misrep-

resent the value of a property for the purpose of influencing a lender to fund a loan for it [32].

Money Laundering. Money laundering is a method used by criminals to insert proceeds obtained from illicit ventures into valid businesses. This conceals the origin of the money, giving them the appearance of legitimate income and making it difficult to track their crimes. Money laundering is also undesirable as it enables the criminals to have economic influence [32].

4 Computational Intelligence and Data Mining Methods for Financial Fraud Detection

A large number of statistical and computational techniques exist that have been applied to data mining problems in recent years (see Figure 8). This section contains a description of the operation of each method used in the reviewed literature, and Table 2 lists the relative strengths and limitations of each.

Bayesian Belief Networks. Bayesian belief networks are a statistical classification technique that makes use of Bayes theorem, a method to determine the probability that a given hypothesis is true. The theorem states that for a hypothesis H (such as whether an object X can be classified within a given class), the probability P is given by:

$$P(H|X) = \frac{(P(X|H) * P(H))}{P(X)} \quad (1)$$

A Bayesian belief network uses a classifier to calculate $P(C_i|X)$ for all possible classes C_i and inserts X into the class with the highest $P(C_i|X)$. In this way the network is shown to classify each sample into the class that it is most likely to belong to [28].

Graphically, a Bayesian belief network can be modelled as a directed acyclic graph, with nodes to represent the samples and edges to reflect a causal dependency between them (see Figure 3). Missing edges can then demonstrate that two variables are independent of one another [32].

Logistic Regression or Logistic Model. Logistic regression is a statistical method of classifying binary data which uses a linear model, often referred to as a logistic or logit model, to perform regression on a set of variables [32], [38]. It is a commonly used method for predicting patterns in data with unambiguous or numeric attributes [32], [7]. Logistic regression makes use of a series of input vectors and a dependant response variable, using the natural logarithm to calculate the probability that the result lies within a particular category. For binary classification the response variable takes the form:

$$y_i = \begin{cases} 1 \\ 0 \end{cases} \quad (2)$$

And the formula for calculating that a sample x_i belongs in class 1 is given by:

$$P(Yi = 1|Xi) = \frac{\exp(w_0 + w^T x_i)}{1 + \exp\,(w_0 + w^T x_i)} \quad (3)$$

Where w_0 and w are the regression tuning parameters representing the intercept and coefficient vector respectively [38].

Neural Network. A neural network is a computational approximation of the human brain which uses a graph of vertices and edges to represent neurons and synapses [32]. The network performs by modelling the input variables as a layer of vertices and then assigning a weighting to each connection in the graph, while the other vertices are put into separate levels reflecting their distance from the input nodes [28] (see Figure 4).

Each node considers its input as a function of the vertices connected to it at the previous layer. For each neuron j the signal received is given by:

$$u_j = \sum w_{ij} * x_i \quad (4)$$

Where w_{ij} is the weight of the link between neurons i and j and x_i is the input of i. If the result is greater than a predetermined amount the current neuron "fires" and becomes an input for the next layer [28].

A back propagation neural network is trained by running samples from a set of training data through the network and comparing the results. For the first iteration the weights at each edge are normally determined randomly, and after the results have been calculated each weighting is adjusted slightly for the next sequence [57]. This continues until either the network has reduced its error to an acceptable amount or a predetermined iteration limit has been reached. After training the network's performance may be tested with a set of validation data [32]. A common problem with back propagation neural networks is overtraining, which can cause the network to focus on tendencies particular to the training set but not the overall problem [57].

Support Vector Machine. Support vector machines are a classification method which work by converting a linear issue into a higher dimensional feature space. This enables complicated, non-linear problems such as financial fraud detection to be solved by linear classification without increasing the computational complexity. The function used to transform the dataset is called the kernel function, which can be con-

sidered as a mapping of points between the input space and a higher dimensional space. The kernel function is defined by:

$$k(x_1, x_2) = \langle \Phi(x_1), \Phi(x_2) \rangle \quad (5)$$

Where $\Phi: X \rightarrow H$ maps points in input space X to higher-dimensional space H. After applying the kernel function to the dataset a hyperplane is used to separate the classes, which takes the form:

$$\langle w, x \rangle + b = 0 \quad (6)$$

The hyperplane is constructed in such a way as to maximise the separation between both classes, which helps to reduce potential errors caused by overtraining (see Figure 5).

The classification for a support vector machine can therefore be defined as:

$$\sum_i \alpha_i y_i k(x_i, x) + b = 0 \quad (7)$$

The choice of which kernel function to use is dependent on the dataset and classification requirements, though there are many commonly used kernel functions such as the Gaussian radial basis function and polynomial function [7].

Genetic Algorithms and Programming. Genetic algorithms use the concept of population evolution to iteratively improve solutions to the problem. It works by randomly creating a starting generation, then continuously reproducing each population using various techniques and choosing the survivors based on their fitness. Reproduction occurs by taking pairs of parents from the current generation and applying crossover on two points, then randomly mutating a single element of the resulting child. The ability of the children is measured using a fitness function, and the result of this determines which parents and children are chosen to represent the next generation.

Measuring the fitness of the children can be as simple as measuring the percentage of samples they classify correctly. The algorithm terminates once a required fitness has been reached, but to avoid infinite looping a limit may be set on the number of iterations that are run (see Figure 6). Genetic algorithms are similar to neural networks in that they require no prior knowledge of the problem domain and are capable of detecting underlying relationships between the samples [38].

Decision Trees, Forests, and CART. Decision trees are a technique that classifies or predicts data using a tree with internal nodes representing binary choices on attributes and branches representing the outcome of that choice [57] (see Figure 7). The nodes are created such that as a sample traverses the tree it is partitioned into subsets until it is eventually sorted into a mutually exclusive subgroup. Decision trees are also known as Classification and Regression Trees (CART) [28].

A decision forest, or random forest, is a collection of decision trees used to avoid the instability and risk of overtraining that can occur with a single tree [7]. Random forests use separate training data between trees and randomly restrict the pool of attributes available when building each internal node [7]. Another method for reducing overfitting in decision trees is pruning, which involves removal of decision nodes without reducing the overall accuracy of the tree [28]. These methods make random forests robust to overtraining and noise, and as each tree is generated independently there is little additional computational complexity [7]. Additionally, the only two parameters that require adjustment are the number of trees and the set of attributes to choose from when building each node, which makes decision forests simple to generate [7].

Group Method of Data Handling. Group method of data handling is an inductive data mining algorithm that calculates optimal solutions through a series of increasingly accurate models. It begins with a simple model, typically a polynomial of the form:

$$z = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_2^2 + w_5 x_1 x_2 \quad \textbf{(8)}$$

After assessing the validity of the model further models are generated by using linear regression to calculate new coefficients for the polynomial. In this way group method of data handling is continually moving toward an optimal approximation. The generated models are functionally similar to a feed forward neural network, with the input to each neuron being derived from a polynomial representation of neurons in the preceding layer and a single neuron in the output layer. After each iteration the success of the model is calculated as the margin of error between actual and expected outputs. If this reaches a predetermined level of correctness the algorithm concludes [38].

Text Mining. Text mining refers to a specific form of data mining which is performed on plain text. Given this broad description there are a variety of different approaches to text mining, though many work by using one or more common preprocessing steps to transform the data into more quantitative samples:

- Filtering of stop words. Common joining words such as “the” or “is” are removed to simplify the dataset.
- Stemming. This involves reducing derived words to their common base, such as removing plurals and tenses.

- Identification of part-of-speech. This recognises the part of a sentence that a word occurs in and can be helpful in reducing error due to homophones and other language nuances.
- Word frequency analysis. The frequency of word i in document j is calculated as (tf_{ij})
- Word combination. Words are further combined into common concepts using a synonym list or more complicated measures.

Once the text has been converted to a more quantitative form a traditional data mining method is applied to actually perform the classification [13], [20]. An alternative to the above approach is to avoid pre-processing and deliberately include all raw data as a method of detecting abnormalities in the text. Measuring factors such as expression, complexity, and specificity within the text can be used to identify subtle differences between samples [26].

Self-Organising Map. Self-organising maps are a form of artificial neural network which consists of a single matrix of neurons. A non-linear algorithm is used to map inputs from a high-dimensional space to the two-dimensional array of neurons. The mapping is designed to model similar input vectors as neurons that are closer together in the resulting matrix, providing a visualisation of the inputs. A distance or neighbourhood function is used to group the nodes, such as the Euclidean distance formula or Gaussian formula [37]. The clustering function applied to each neutron is given by:

$$Y_{i+1} = Y_i + \alpha(X_i - Y_{i-1}) \quad (\mathbf{9})$$

Where Y_i is the current weighting of a specific node, X_i is the current input vector, and α is the chosen distance function. The clustering step is performed a set number of iterations before the algorithm completes [33].

Process Mining. Process mining involves analysis of transactions and event logs to construct models representing the behaviour of a system. A specific process instance is created to represent the individual cases within the system, and assumptions are applied to the available data to determine whether it is suitable for observation with this particular process. Given a model of expected behaviour within the system, the typical operation steps are:

- Log preparation and inspection. The logs are obtained and pre-processed to remove extraneous noise. If no existing model was available one can be created based on the provided logs.
- Analysis. Models are observed to check various behaviours. Aspects such as control flow, performance, and user-roles can be analysed to determine the expected outcomes of the system.

- Verification. Given the results of the above analysis, the process miner is applied to various samples and determines whether they represent typical system behaviour.

In addition to classification, the ability to generate its own model makes process mining useful for discovering flows and practices within a complicated system [27].

Artificial Immune System. Artificial immune systems are a data mining approach that work by imitating the behaviour of a biological immune system to detect antigens [52]. A variety of biological characteristics can be simulated by the artificial immune system, but most model the creation of detector cells and their ability to detect foreign bodies. The detector cells are generated randomly and the simulation is performed to test and evaluate their effectiveness, similar to the training performed by other classification methods.

One common form is clonal selection, which continually generates detector cells that only live a short time. If a cell detects an antigen it extends its life to fight the intruder, and may also mutate as a result of the conflict. The surviving cells at the end of the simulation are therefore the ones best suited to detecting the antigens. Another common implementation is negative selection, which works by randomly creating cells and determining which of these react with other non-invasive cells within the system. Any that do are discarded, resulting in the remainder being proficient at detecting intruders [43].

Hybrid Methods. Hybrid methods are a combination of multiple traditional methods by selecting beneficial attributes of each in an attempt to create a superior algorithm for a specific problem domain. Hybrid methods can be constructed in a variety of different ways: at the highest level methods can simply be applied linearly, with the outputs of the first providing the inputs of the second [35]. Similarly one method may be used as a pre-processing step to modify the data in preparation for classification [27], or at a lower level the individual steps of the algorithms can be intertwined to create something fundamentally original [16]. Additionally, hybrid methods can be used to tailor solutions to an individual problem domain. Different aspects of performance can be specifically targeted, including classification ability, ease of use, and computational efficiency.

5 Classification of Existing Financial Fraud Detection

In the following section we will classify the financial fraud detection techniques reviewed based on success rate, method chosen, and fraud type studied. This categorisation will enable us to demonstrate trends in current research methods, including which have been successful and any factors that have not been covered.

5.1 Classification Based on Performance

A variety of standards have been used to determine performance, but the three most commonly used are accuracy, sensitivity, and specificity. Accuracy measures the ratio of all successfully classified samples to unsuccessful ones. Sensitivity compares the amount of items correctly identified as fraud to the amount incorrectly listed as fraud, also known as the ratio of true positives to false positives. Specificity refers to the same concept with legitimate transactions, or the comparison of true negatives to false negatives [7], [38].

Tables 3, 4, and 5 classify financial fraud detection research based on these performance measures. Additionally, figure 9 depicts the amount of studies that have been performed on each combination of detection method and fraud type, and the comparative performance of each

In addition to the three performance measures discussed here, several other performance measures have been used in the literature. For example, Duman et al. chose to show their results for sensitivity in graph form instead of deterministic values, grouped by each set of input parameters [16]. In addition to other forms of graphing [37], some research used software-determined success levels or case-based procedures to determine the success of their fraud detection techniques [41], [27].

From the results we can see that CI methods typically had better success rate than statistical methods. Sensitivity was slightly better for random forests and support vector machines than logistic regression, with comparable specificity and accuracy [7]. Genetic programming, support vector machines, probabilistic neural networks, and group method of data handling outperformed regression in all three areas [38]. Additionally, a neural network with exhaustive pruning was found to be more specific and accurate than CDA [10]. One statistical method seems to contradict this theory however: Bayesian belief networks were reported to be more accurate than neural networks and decision trees [28], and adding Bayesian logic to regression outperformed logistic regression alone [6].

Most of the research showed a large difference between each method's sensitivity and specificity results. For example, Bhattacharyya et al. showed that logistic regression, support vector machines and random forests all performed significantly better at detecting legitimate transactions correctly than fraudulent ones [7]. Support vector machines, genetic programming, neural networks, group method of data handling, and particularly logistic regression were also slightly less sensitive [38]. And a neural network with exhaustive pruning showed more specificity than sensitivity [10].

As explained previously, fraud detection is a problem with a large difference in misclassification costs: it is typically far more expensive to misdiagnose a fraudulent transaction as legitimate than the reverse. With that in mind it would be beneficial for a detection techniques to show a much higher sensitivity than specificity, meaning that these results are less than ideal. Contrary to this belief, Hoogs et al. hypothesised that financial statement fraud may carry higher costs for false positives, and their results reflect this with a much higher specificity [24]. Panigrahi et al. also acknowledged the costs associated with following up credit card transactions marked as fraudulent, focussing their results on sensitivity only [35]. The CDA and CART

methods, as well as the neural networks, Bayesian belief networks and decision trees performed better in this regard, with all showing a somewhat higher ability to classify fraudulent transactions than legitimate ones [10], [28].

5.2 Classification Based on Detection Algorithm

Classifying fraud detection practices by the detection algorithm is a useful way to identify the suitable techniques for this problem domain. It can also help us to determine why particular methods were chosen or successful. Additionally, we can identify any gaps in research by looking at algorithms which have not been explored sufficiently. Table 6 shows classification of financial fraud detection practices based on detection algorithm (conventional data mining and CI-based approaches) used.

Previously it was mentioned that early fraud detection research focussed on statistical models and neural networks; however, it may be noted that these methods still continue to be popular. Many used at least one form of neural network [28], [38], [10], some investigated logistic regression [7], [36], [49], [38], [25], while others applied Bayesian belief networks and other Bayesian algorithms [23], [28], [6]. Application of CDA has been relatively uncommon [10]. Neural networks and logistic regression are often chosen for their well-established popularity, giving them the ability to be used as a control method by which other techniques are tested. Comparatively, more advanced methods such as support vector machines and genetic programming have received substantially less attention. Yue et al. also reported that all the methods mentioned in their research were a form of classification, with no studies performed on clustering or time-series approaches, and that most of the research focussed on supervised learning as opposed to unsupervised [55].

Several of the research focussed on a single form of fraud detection which they advocated above others, such as studying text mining with the singular validation decomposition vector [20], self-organising maps [37], [33], logistic regression [49], [36], and fuzzy logic [41]. Additionally, some researchers focussed soley on classification and regression trees [3],[40], Bayesian belief networks [23], individual statistical techniques [35], artificial immune systems [51], [43] or their own hybrid methods [16]. This unilateral approach is useful for demonstrating the ability of the specific method in isolation, but without comparing it to other methods it is difficult to understand the relative performance of the technique. Additional factors such as the fraud type researched and the specific dataset used can influence the results of the experiment. Future research could focus on reviewing these methods as against other more established techniques.

A rising trend in fraud detection is the use of hybrid methods which utilise the strengths of multiple algorithms to classify samples. Duman and Ozcelik used a combination of scatter search and genetic algorithm, based on the latter but targeting attributes of scatter search such as the smaller populations and recombination as the reproduction method [16]. A different approach was taken by Panigrahi et al. who used two methods sequentially, beginning with the Depster-Schaefer method to combine rules and then using a Bayesian learner to detect the existence of fraud [35]. Some researchers applied fuzzy logic to introduce variation to their samples,

attempting to transform it to resemble real world data before deploying a different technique to actually detect the presence of fraud [27]. The investigators recognised that applying 'fuzziness' to their problem increased the performance of their solution [52]. Similarly, several researchers combined traditional computational intelligence methods with text mining to analyse financial statements for the presence of fraud [13], [26], [15].

5.3 Classification Based on Fraud Type

Given the varying nature of each type of fraud, the problem domain can differ significantly depending on the form that is being detected. By classifying the existing practices on the type of fraud investigated we can identify the techniques more suitable and more commonly used for a specific type of fraud. Additionally we can infer the varieties which are considered the most important for investigation depending on the scope and scale of their impact. Table 7 depicts the classification based on fraud types considered, along with the detection methods used. Table 9 lists the datasets used by each researcher.

As with each chosen algorithm, feature selection will differ depending on the problem domain. Specific financial statement fraud exists within individual companies, and as such attribute ratios are used instead of absolute values. Koh and Low provide a good example of the relevant ratios such as net income to total assets, interest payments to earnings before interest and tax, and market value of equity to total assets [29]. In comparison, research into credit card fraud has typically selected independent variables or aggregate values which may be quantitative or qualitative. For example, Bhattacharyya et al. made use of transaction amount, categorical values such as account number, transaction date, and currency, and aggregated properties like total transaction amount per day, and average amount spent at a single merchant [7].

We can see that the existing research has been greatly unbalanced in fraud type studied. The vast majority of papers have focussed on two forms of financial fraud: *credit card fraud* and *financial statement fraud*. Only a handful of studies have looked at securities and commodities fraud, and many focussed on external forms of corporate fraud while neglecting the internal ones [27]. Ngai et al. found that insurance fraud had received the highest coverage during their research [32]: the fact that we identified only a few examples of published literature on this type of fraud since 2007 indicates that research into insurance fraud is declining. Additionally, no studies have been performed directly on mortgage fraud or money laundering. Reasons for this disparity may include differing relevance to stakeholders of each fraud type.

To determine which method was the most successful for each fraud type we looked at the accuracy scores reported in the experimental research. There are many metrices used to assess performance but accuracy was the most commonly used and therefore provided the best basis for comparison. Analysis of the individual problem domain is required to completely assess the usefulness of each method, but the results shown

here should provide a useful starting point for future researchers to investigate new fraud detection algorithms. It should be noted that we have reported only the best accuracy result obtained by each method for specific fraud types; the test conditions for individual experiments are not necessarily comparable. Table 8 shows these results for each fraud type, as well as comparable methods.

All the solutions used for credit card fraud had a very high success rate, including regression, support vector machines, artificial immune systems, and random forests. Out of these the self organising map is recommended due to its perfect accuracy with a 10000 transaction sample [33]. As mentioned earlier there has been a lack of recent research into insurance fraud, with only one researcher comparing typical logistic regression to a hybrid version with Bayesian logic. In this case the hybrid method offered a significant improvement with an accuracy of 99.5% compared to 60.7% [6]. Financial statement fraud had a large variance in results, from CDA at 71.4% to a probibalistic neural network with 98.1% [10], [38]. Several other methods also had success rate greater than 90%, including Bayesian belief networks, support vector machine, genetic programming, group method of data handling, and some hybrid methods based on text mining.

6 Financial Fraud Detection Challenges and Future Directions

Financial fraud detection is an evolving field in which it is desirable to stay ahead of the perpetrators. Additionally, it is evident that there are still facets of intelligent fraud detection that have not been investigated. In this section we present some of the key issues associated with financial fraud detection and suggest areas for future research. Some of the identified issues and challenges are as follows:

- *Typical classification problems:* CI and data mining-based financial fraud detection is subject to the same issues as other classification problems, such as feature selection, parameter tuning, and analysis of the problem domain.
- *Fraud types and detection methods:* Financial fraud is a diverse field and there has been a large imbalance in both fraud types and detection methods studied: some have been studied extensively while others, such as hybrid methods, have only been looked at superficially.
- *Privacy considerations:* Financial fraud is a sensitive topic and stakeholders are reluctant to share information on the subject. This has led to experimental issues such as undersampling.
- *Computational performance:* As a high-cost problem it is desirable for financial fraud to be detected immediately. Very little research has been conducted on the computational performance of fraud detection methods for use in real-time situations.
- *Evolving problem:* Fraudsters are continually modifying their techniques to remain undetected. As such detection methods are required to be able to constantly adapt to new fraud techniques.

- *Disproportionate misclassification costs:* Fraud detection is primarily a classification problem with a vast difference in misclassification costs. Research on the performance of detection methods with respect to this factor is an area which needs further attention.
- *Generic framework:* Given that there are many varieties of fraud, a generic framework which can be applied to multiple fraud categories would be valuable.

As a classification problem, financial fraud detection suffers from the same issues as other similar problems. Feature selection has a high impact on the success of any classification method. While some researchers have mentioned feature selection for one type of fraud [29], [7], no comparisons have been made between features for differing problem domains. Also, one of the major benefits of the computational intelligence and data mining methods is their ability to be adjusted to fit the problem domain. Existing research has rarely used any form of customisation or tuning for specific problems; however, tuning is an important factor in the context of an algorithm's performance. For example, the number of nodes and internal layers within a neural network has a large impact on both accuracy and computational performance. Similarly the kernel function chosen will considerably alter the success of a support vector machine and parameters such as the fitness function, crossover method, and probability for mutation will impact the results of a genetic programming algorithm. Research on customisation or tuning of the computational methods is required to truly comprehend the ability of each method. Further, in other data mining cases the solution algorithm is selected based on its performance within the problem domain, which for financial fraud detection is the type of fraud investigated. Studies on the suitability of various methods for each fraud category are necessary to understand which attributes of each algorithm make them appropriate for detecting financial fraud.

From the existing literature it is apparent that there are some forms of fraud that have not been investigated as extensively as others. Financial statement fraud has been considerably investigated, which is understandable given its high profile nature, but there are other forms of fraud that have a significant impact on consumers. Credit card fraud often has a direct impact on the public and the recent increase in online transactions has led to a majority of the U.S. public being concerned with identity theft [7]. A benefit of this close relation to the user is that credit card fraud is typically detected quickly, which gives researchers access to large datasets of unambiguous transactions. Other forms of fraud which have not been covered in depth include money laundering, mortgage, and securities and commodities fraud. A lack of sufficient sample size may be the reason for the lack of research in these areas [32]. Future studies that focussed on these types of fraud detection would be beneficial.

The private nature of financial data has led to institutions being reluctant to share fraudulent information. This has had an affect both on the fraud types that have been investigated as well as the datasets used for the purpose. In the published literature many of the financial fraud simulations consisted of less than a few hundred samples, typically with comparable amounts of fraudulent and legitimate specimens. This is

contrary to the realities of the problem domain, where fraud cases are far outweighed by legitimate transactions [7]. Undersampling the problem domain like this can cause biases in the data that do not accurately represent real-world scenarios [24]. There is a definite need for further studies with realistic samples to accurately depict the performance of each method [20].

Some forms of financial fraud occur very rapidly, such as credit card fraud. If a fraudster obtains an individual's credit card information it's very likely that they will use it immediately until the card limit is reached. The ability to detect fraud in real-time would be highly beneficial as it may be able to prevent the fraudster from making subsequent transactions. Computational performance is therefore a key factor to consider in fraud detection. Though some researchers have noted the performance of their particular methods [7], [37], most studies were simulations performed on test datasets. Further research focussing on the computational as well as classification performance is required.

Unlike many classification problems, fraud detection solutions must be capable of handling active attempts to circumvent them. As detection methods become more intelligent, fraudsters are also constantly upgrading their techniques. For example, in the last few decades credit card fraud has moved from individuals stealing or forging single cards to large-scale phone and online fraud perpetrated by organised groups [7]. It is therefore necessary for fraud detection methods to be capable of evolving to stay ahead of fraudsters. Some researchers have considered models for adaptive classification, however further research is required to fully develop these for use in practical fraud detection problems [59].

As explained previously fraud has a large cost to businesses. Additionally, fraud detection has associated costs: systems require maintenance and computational power, and auditors must be employed to monitor them and investigate when a potential fraud case is identified [28]. The expense of a false positive, in misclassifying a legitimate transaction as fraud, is typically far less than that of a false negative [32]. Insufficient study has been performed on the disproportionate nature of these costs, with attention typically focussing on the traditional classification performance methods outlined in Section 5.1. Considering the accuracy of each fraud detection method, focus should be on achieving an optimum balance for each technique such that the expense is smallest. Research specifically focused on finding this balance would add significant real-world value to financial fraud detection.

Given the diversity of common categories of fraud it would be useful to have some form of generic framework that could apply to more than one fraud category. Such a framework could be used to study the differences between various types of fraud, or even specific details such as differentiating between stolen and counterfeit credit cards [7]. A ubiquitous model could also be used to determine which specific fraud detection method is applicable given the problem domain. This approach has been investigated slightly with response surface methodology [59], but more detailed research is desirable.

7 Conclusion

Fraud detection is an important part of the modern finance industry. This literature review studied research into intelligent approaches to fraud detection, both statistical and computational. Though their performance differed, each technique was shown to be reasonably capable at detecting various forms of financial fraud. In particular, the ability of the computational methods such as neural networks and support vector machines to learn and adapt to new techniques is highly effective to the evolving tactics of fraudsters.

There are still many aspects of intelligent fraud detection that have not yet been the subject of research. Some types of fraud, as well as some data mining methods, have been superficially explored but require future study to be completely understood. There is also the opportunity to examine the performance of existing methods by adjusting their parameters, as well as the potential to study cost benefit analysis of computational fraud detection. Finally, further research into the differences between each type of financial fraud could lead to a general framework which would greatly improve the accuracy of intelligent detection methods.

References

1. Abbott LJ, Park Y, and Parker S (2000) The effects of audit committee activity and independence on corporate fraud. *Managerial Finance* **26**, 55-68.
2. Aleskerov E, Freisleben B, and Rao B (1997) Cardwatch: A neural network based database mining system for credit card fraud detection. In *Computational Intelligence for Financial Engineering (CIFEr), 1997., Proceedings of the IEEE/IAFE 1997*. (ed.), Vol. pp. 220-6, IEEE,
3. Bai B, Yen J, and Yang X (2008) False financial statements: characteristics of China's listed companies and CART detecting approach. *International Journal of Information Technology & Decision Making* **7**, 339-59.
4. Bell TB and Carcello JV (2000) A decision aid for assessing the likelihood of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory* **19**, 169-84.
5. Beneish MD (1999) The detection of earnings manipulation. *Financial Analysts Journal* 24-36.
6. Bermúdez L, Pérez J, Ayuso M, Gómez E, and Vázquez F (2008) A Bayesian dichotomous model with asymmetric link for fraud in insurance. *Insurance: Mathematics and Economics* **42**, 779-86.
7. Bhattacharyya S, Jha S, Tharakunnel K, and Westland JC (2011) Data mining for credit card fraud: A comparative study. *Decision Support Systems* **50**, 602-13.
8. Bolton RJ and Hand DJ (2002) Statistical fraud detection: A review. *Statistical Science* 235-49.
9. Bolton RJ and Hand DJ (2001) Unsupervised profiling methods for fraud detection. *Credit Scoring and Credit Control VII* 235-55.
10. Bose I and Wang J (2007) Data mining for detection of financial statement fraud in Chinese Companies. Paper presented at the International Conference on Electronic Commerce, Administration, Society and Education, Hong Kong, 15-17 August 2007.

11. Busta B and Weinberg R (1998) Using Benford's Law and neural networks as a review procedure. *Managerial Auditing Journal* **13**, 356-66.
12. Calderon TG and Cheh JJ (2002) A roadmap for future neural networks research in auditing and risk assessment. *International Journal of Accounting Information Systems* **3**, 203-36.
13. Cecchini M, Aytug H, Koehler GJ, and Pathak P (2010) Making words work: Using financial text as a predictor of financial events. *Decision Support Systems* **50**, 164-75.
14. Deshmukh A and Talluru L (1998) A rule-based fuzzy reasoning system for assessing the risk of management fraud. *International Journal of Intelligent Systems in Accounting, Finance & Management* **7**, 223-41.
15. Dong W, Liao SS, Fang B, Cheng X, Chen Z, and Fan W (2014) THE DETECTION OF FRAUDULENT FINANCIAL STATEMENTS: AN INTEGRATED LANGUAGE MODEL.
16. Duman E and Ozcelik MH (2011) Detecting credit card fraud by genetic algorithm and scatter search. *Expert Systems with Applications* **38**, 13057-63.
17. Fanning KM and Cogger KO (1998) Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems in Accounting, Finance & Management* **7**, 21-41.
18. Feroz EH, Kwon TM, Pastena VS, and Park K (2000) The efficacy of red flags in predicting the SECs targets: an artificial neural networks approach. *International Journal of Intelligent Systems in Accounting, Finance & Management* **9**, 145-57.
19. Fraser IA, Hatherly DJ, and Lin KZ (1997) AN EMPIRICAL INVESTIGATIONOF THE USE OF ANALYTICAL REVIEW BY EXTERNAL AUDITORS. *The British Accounting Review* **29**, 35-47.
20. Glancy FH and Yadav SB (2011) A computational model for financial reporting fraud detection. *Decision Support Systems* **50**, 595-601.
21. Green BP and Choi JH (1997) Assessing the risk of management fraud through neural network technology. *Auditing-A Journal Of Practice & Theory* **16**, 14-28.
22. Hansen J, McDonald JB, Messier Jr W, and Bell TB (1996) A generalized qualitative-response model and the analysis of management fraud. *Management Science* **42**, 1022-32.
23. Holton C (2009) Identifying disgruntled employee systems fraud risk through text mining: A simple solution for a multi-billion dollar problem. *Decision Support Systems* **46**, 853-64.
24. Hoogs B, Kiehl T, Lacomb C, and Senturk D (2007) A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud. *Intelligent Systems in Accounting, Finance and Management* **15**, 41-56.
25. Huang SY (2013) Fraud Detection Model by Using Support Vector Machine Techniques. *JDCTA: International Journal of Digital Content Technology and its Applications* **7**, 32-42.
26. Humpherys SL, Moffitt KC, Burns MB, Burgoon JK, and Felix WF (2011) Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems* **50**, 585-94.
27. Jans M, van der Werf JM, Lybaert N, and Vanhoof K (2011) A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications* **38**, 13351-9.
28. Kirkos E, Spathis C, and Manolopoulos Y (2007) Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications* **32**, 995-1003.
29. Koh HC and Low CK (2004) Going concern prediction using data mining techniques. *Managerial Auditing Journal* **19**, 462-76.

30. Kou Y, Lu C-T, Sirwongwattana S, and Huang Y-P (2004) Survey of fraud detection techniques. In *Networking, sensing and control, 2004 IEEE international conference on*. (ed.), Vol. 2, pp. 749-54, IEEE,
31. Maes S, Tuyls K, Vanschoenwinkel B, and Manderick B (2002) Credit card fraud detection using Bayesian and neural networks. In *Proceedings of the 1st international naiso congress on neuro fuzzy technologies*. (ed.), Vol. pp.
32. Ngai E, Hu Y, Wong Y, Chen Y, and Sun X (2011) The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems* **50**, 559-69.
33. Olszewski D (2014) Fraud detection using self-organizing map visualizing the user profiles. *Knowledge-Based Systems*
34. Paasch CA (2010) In *Credit card fraud detection using artificial neural networks tuned by genetic algorithms*. Vol. pp. HONG KONG UNIV. OF SCI. AND TECH.(HONG KONG),
35. Panigrahi S, Kundu A, Sural S, and Majumdar AK (2009) Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. *Information Fusion* **10**, 354-63.
36. Pinquet J, Ayuso M, and Guillen M (2007) Selection bias and auditing policies for insurance claims. *Journal of Risk and Insurance* **74**, 425-40.
37. Quah JT and Sriganesh M (2008) Real-time credit card fraud detection using computational intelligence. *Expert Systems with Applications* **35**, 1721-32.
38. Ravisankar P, Ravi V, Raghava Rao G, and Bose I (2011) Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems* **50**, 491-500.
39. Rezaee Z (2002) In *Financial statement fraud: prevention and detection*. Vol. pp. John Wiley & Sons,
40. Sahin Y, Bulkan S, and Duman E (2013) A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications* **40**, 5916-23.
41. Sánchez D, Vila M, Cerda L, and Serrano J-M (2009) Association rules applied to credit card fraud detection. *Expert Systems with Applications* **36**, 3630-40.
42. Sohl JE and Venkatachalam A (1995) A neural network approach to forecasting model selection. *Information & Management* **29**, 297-303.
43. Soltani Halvaiee N and Akbari MK (2014) A novel model for credit card fraud detection using Artificial Immune Systems. *Applied Soft Computing*
44. Spathis C, Doumpos M, and Zopounidis C (2002) Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques. *European Accounting Review* **11**, 509-35.
45. Spathis CT (2002) Detecting false financial statements using published data: some evidence from Greece. *Managerial Auditing Journal* **17**, 179-91.
46. Stolfo S, Fan W, Lee W, Prodromidis A, and Chan P (1997) Credit card fraud detection using meta-learning: Issues and initial results. In *AAAI-97 Workshop on Fraud Detection and Risk Management*. (ed.), Vol. pp.
47. Summers SL and Sweeney JT (1998) Fraudulently misstated financial statements and insider trading: an empirical analysis. *Accounting Review* 131-46.
48. Vatsa V, Sural S, and Majumdar AK (2005) A game-theoretic approach to credit card fraud detection. In *Information Systems Security*. Vol. pp. 263-76. Springer,
49. Viaene S, Ayuso M, Guillen M, Van Gheel D, and Dedene G (2007) Strategies for detecting fraudulent claims in the automobile insurance industry. *European Journal of Operational Research* **176**, 565-83.

50. Whitrow C, Hand DJ, Juszczak P, Weston D, and Adams NM (2009) Transaction aggregation as a strategy for credit card fraud detection. *Data Mining and Knowledge Discovery* **18**, 30-55.
51. Wong N, Ray P, Stephens G, and Lewis L (2012) Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results. *Information Systems Journal* **22**, 53-76.
52. Wu SX and Banzhaf W (2008) Combatting financial fraud: a coevolutionary anomaly detection approach. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*. (ed.), Vol. pp. 1673-80, ACM,
53. Yang W-S and Hwang S-Y (2006) A process-mining framework for the detection of healthcare fraud and abuse. *Expert Systems with Applications* **31**, 56-68.
54. Yeh I and Lien C-h (2009) The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications* **36**, 2473-80.
55. Yue D, Wu X, Wang Y, Li Y, and Chu C-H (2007) A review of data mining-based financial fraud detection research. In *Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on*. (ed.), Vol. pp. 5519-22, IEEE,
56. Zaki M and Theodoulidis B (2013) Analyzing Financial Fraud Cases Using a Linguistics-Based Text Mining Approach. *Available at SSRN 2353834*
57. Zhang D and Zhou L (2004) Discovering golden nuggets: data mining in financial application. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* **34**, 513-22.
58. Zhang G, Eddy Patuwo B, and Y Hu M (1998) Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting* **14**, 35-62.
59. Zhou W and Kapoor G (2011) Detecting evolutionary financial statement fraud. *Decision Support Systems* **50**, 570-5.
60. West J, Bhattacharya M and Islam R (2014) "Intelligent Financial Fraud Detection Practices: An Investigation", in Proceedings of the 10th International Conference on Security and Privacy in Communication Networks (SecureComm 2014).
61. FBI, Federal Bureau of Investigation, Financial Crimes Report to the Public Fiscal Year, Department of Justice, United States, 2010-2011, http://www.fbi.gov/stats-services/publications/financial-crimes-report-2010-2011.
62. Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.
63. Edge, M. E., & Sampaio, P. R. F. (2009). A survey of signature based methods for financial fraud detection. *computers & security*, *28*(6), 381-394.

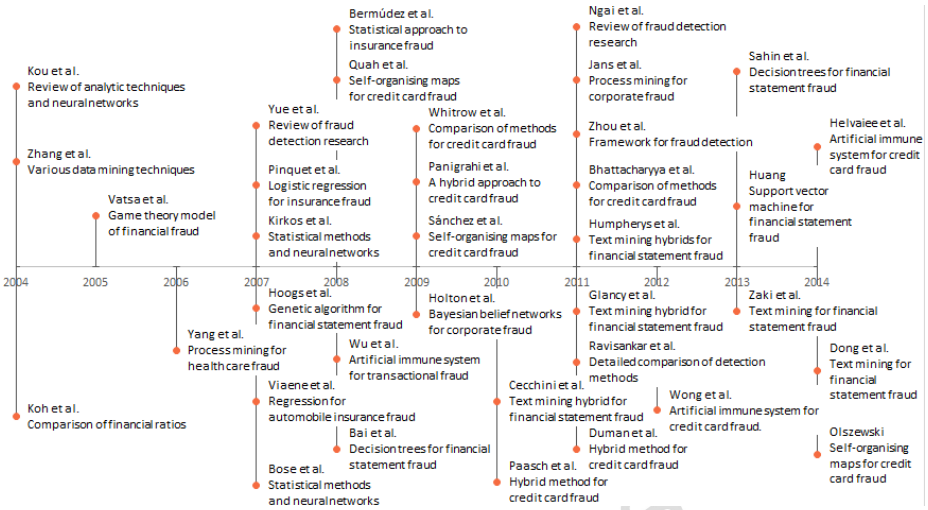

Fig. 1. Recent financial fraud detection research

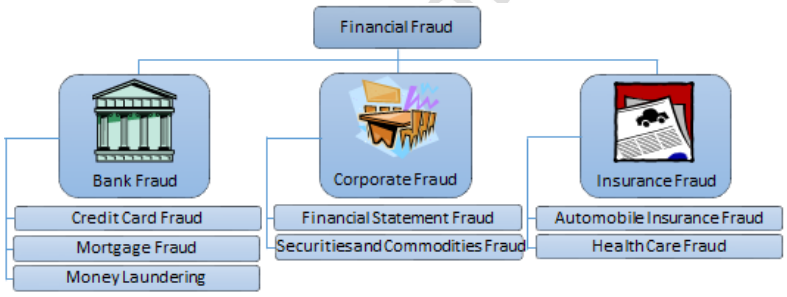

Fig. 2. Common types of financial fraud.

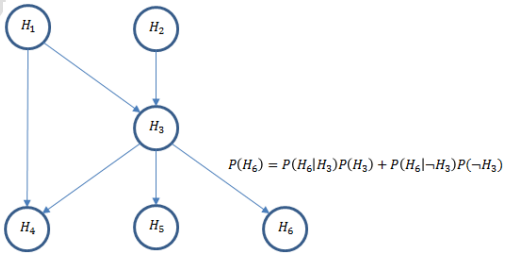

Fig. 3. Example graphical representation of a Bayesian belief network showing the causal relationship between hypothesis H_6 and H_3.

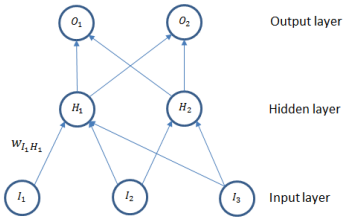

Fig. 4. A simple neural network with three input variables, two outputs, and a single hidden layer.

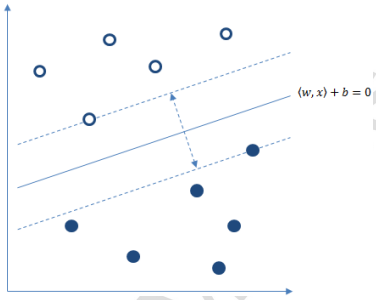

Fig. 5. Two dimensional example showing the margin of separation between support vectors and hyperplane.

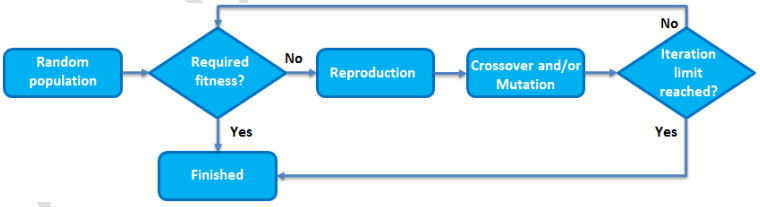

Fig. 6. Standard genetic algorithm process.

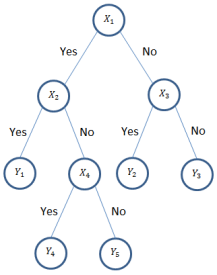

Fig. 7. Binary decision tree where the input variables X_i will be a list of input predicates and the outputs Y_i will classify the results into one of multiple classes.

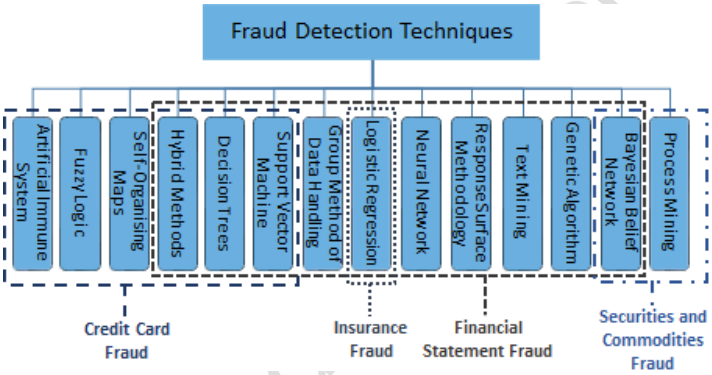

Fig. 8. Detection methods and fraud types.

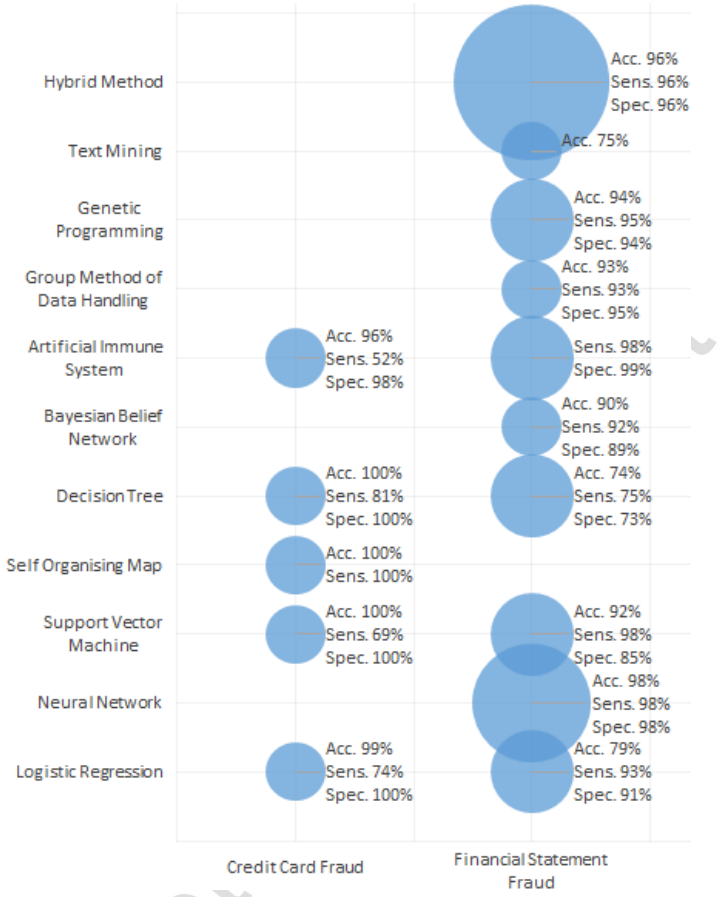

Fig. 9. Comparative performance of various detection methods.

Table 1. Examples of research on fraud detection prior to 2004

Method investigated	Year of publication	Research
Neural network	1997	[2], [21]
	1998	[11]
	2000	[18]
	2002	[12]
Regression	1998	[47]
	1999	[5]
	2000	[1], [4]
	2002	[44], [45]
Fuzzy logic	1998	[14]

Other statistical methods	1996	[22]
	1997	[46]

Table 2. Relevant properties of data mining methods for financial fraud detection

Method	Strengths	Limitations
Neural network	Well established history with fraud detection. Proven suitability with other non-algorithmic, binary classification problems.	Requires high computational power for training and operation, making it unsuitable for real-time function. Potential for overfitting if training set is not a good representation of the problem domain, so requires constant retraining to adapt to new methods of fraud.
Logistic model	Simple to implement. Well established history with fraud detection.	Lower classification performance than other data mining methods, difficulty with the complexity of fraud detection
Support vector machine	Capable of solving non-linear classification problems like fraud detection. Training and operation requires low computational power, which gives potential for real-time operation.	Difficult for auditors to process results due to transformation of input set
Decision trees, forests and CART	Simple to implement and understand. Training and operation requires low computational power, which gives potential for real-time operation.	Potential for overfitting if training set is not a good representation of the problem domain, so requires constant retraining to adapt to new methods of fraud. Optimisation during initial setup requires high computational power.
Genetic algorithm/programming	Simple to implement using classification accuracy as the fitness solution. Proven suitability with other non-algorithmic, binary classification problems.	Requires high computational power for training and operation, making it unsuitable for real-time function. Difficulty adapting to new fraud methods due to local maxima/minima problem.
Text mining	Highly useful for fraud types with large amounts of textual data, such as financial statement fraud.	Requires another classification method to perform the actual fraud detection. Textual data is typically more

		subjective, and thus harder to process.
Group method of data handling	Simple to implement. Guaranteed to provide the best available solution.	Difficulty classifying noisy data, which many fraud types contain.
Response-surface methodology	Capable of solving non-linear classification problems like fraud detection.	Lower classification performance than other data mining methods, difficulty with complexity of fraud detection.
Self-organizing map	Simple to implement and very easy for auditors to understand given visual nature of results.	Visualisation requires auditor observation, cannot be fully automated easily.
Bayesian belief network	Proven suitability with other non-algorithmic, binary classification problems. High computational efficiency gives potential for real-time operation.	Requires strong understanding of typical and abnormal behaviour for the investigated fraud type.
Process mining	Useful for internal fraud investigations where information is available for every iterative step. Ability to focus on an entire process chain instead of individual attributes.	Requires strong understanding of typical and abnormal behaviour for the investigated fraud type. Difficulty classifying noisy data, which many fraud types contain.
Artificial immune system	High suitability for classification problems with imbalanced data, such as fraud detection.	Requires high computational power for operation, making it unsuitable for real-time function.
Hybrid methods	Adaptive to new fraud techniques by combining the strengths of multiple traditional detection methods.	Fraud is a high-cost problem and therefore new, under-tested methods carry a large amount of risk.

Table 3. Accuracy results.

Research	Fraud Investigated	Method Investigated	Accuracy
[7]	Credit card transaction fraud from a real world example	Logistic model (regression)	96.6-99.4%
		Support vector machines	95.5-99.6%
		Random forests	97.8-99.6%
[33]	Credit card transaction fraud from a bank in	Self-organising map	100%

	Warsaw		
[43]	Credit card fraud from a Brazilian bank	Artificial immune system	94.6-96.4%
[6]	Insurance fraud from automobile insurance claims for a Spanish company	Logistic regression	60.680%
		Bayesian skewed regression	99.538%
[28]	Financial statement fraud from a selection of Greek manufacturing firms	Decision trees	73.6%
		Neural networks	80%
		Bayesian belief networks	90.3%
[38]	Financial statement fraud with financial items from a selection of public Chinese companies	Support vector machine	70.41-73.41%
		Genetic programming	89.27-94.14%
		Neural network (feed forward)	75.32-78.77%
		Group method of data handling	88.14-93.00%
		Logistic model (regression)	66.86-70.86%
		Neural network (probabilistic)	95.64-98.09%
[20]	Financial statement fraud with managerial statements for US companies	Text mining with singular validation decomposition vector	95.65%
[13]	Financial statement fraud with managerial statements for US companies	Text mining	45.08-75.41%
		Text mining and support vector machine hybrid	50.00-81.97%
[26]	Financial statement fraud with managerial statements for US companies	Text mining and decision tree hybrid	67.3%
		Text mining and Bayesian belief network hybrid	67.3%
		Text mining and support vector machine hybrid	65.8%
[25]	Financial statement fraud from Taiwanese lawsuits	Logistic regression	19%-79%
		Support vector machine	71%-92%
[10]	Financial statement fraud with financial items from a selection	CDA	71.37%

	of public Chinese companies		
		CART	72.38%
		Neural network (exhaustive pruning)	77.14%

Table 4. Sensitivity results.

Research	Fraud Investigated	Method Investigated	Sensitivity
[7]	Credit card transaction fraud from a real world example	Logistic model (regression)	24.6-74.0%
		Support vector machines	43.0-68.7%
		Random forests	42.3-81.2%
[33]	Credit card transaction fraud from a bank in Warsaw	Self-organising map	100%
[43]	Credit card fraud from a Brazilian bank	Artificial immune system	33.6%-52.6%
[6]	Insurance fraud from automobile insurance claims for a Spanish company	Logistic regression	85.149%
		Bayesian skewed regression	85.149%
[28]	Financial statement fraud from a selection of Greek manufacturing firms	Decision trees	75.0%
		Neural networks	82.5%
		Bayesian belief networks	91.7%
[38]	Financial statement fraud with financial items from a selection of public Chinese companies	Support vector machine	55.43-73.60%
		Genetic programming	85.64-95.09%
		Neural network (feed forward)	67.24-80.21%
		Group method of data handling	87.44-93.46%
		Logistic model (regression)	62.91-65.23%
		Neural network (probabilistic)	87.53-98.09%
[20]	Financial statement fraud with managerial statements	Text mining with singular validation decomposition vector	95.65%
[25]	Financial statement fraud from Taiwanese lawsuits	Logistic regression	24-93%
		Support vector machine	76-98%

[10]	Financial statement fraud with financial items from a selection of public Chinese companies	CDA	61.96%
		CART	72.40%
		Neural network (exhaustive pruning)	80.83%
[35]	Credit card fraud using legitimate customer transaction history as well as generic fraud transactions	Bayesian learning with Dempster-Shafer combination	71-83%
[24]	Financial statement fraud from Accounting and Auditing Enforcement Releases by the Securities and Exchange Commission	Genetic algorithm	13-27%
[52]	Transactional fraud in automated bank machines and point of sale from a financial institution	Coevolution artificial immune system	97.688-98.266%
		Standard evolution artificial immune system	92.486-95.376%

Table 5. Specificity results.

Research	Fraud Investigated	Method Investigated	Specificity
[7]	Credit card transaction fraud from a real world example	Logistic model (regression)	96.7-99.8%
		Support vector machines	95.7-99.8%
		Random forests	97.9-99.8%
[43]	Credit card fraud from a Brazilian bank	Artificial immune system	97.8-98.1%
[6]	Insurance fraud from automobile insurance claims for a Spanish company	Logistic regression	60.430%
		Bayesian skewed regression	99.677%
[28]	Financial statement fraud from a selection of Greek manufactur-	Decision trees	72.5%

	ing firms		
		Neural networks	77.5%
		Bayesian belief networks	88.9%
[38]	Financial statement fraud with financial items from a selection of public Chinese companies	Support vector machine	70.41-73.41%
		Genetic programming	89.27-94.14%
		Neural network (feed forward)	75.32-78.77%
		Group method of data handling	88.34-95.18%
		Logistic model (regression)	70.66-78.88%
		Neural network (probabilistic)	94.07-98.09%
[20]	Financial statement fraud with managerial statements	Text mining with singular validation decomposition vector	95.65%
[25]	Financial statement fraud from Taiwanese lawsuits	Logistic regression	24-91%
		Support vector machine	11-85%
[10]	Financial statement fraud with financial items from a selection of public Chinese companies	CDA	80.77%
		CART	72.36%
		Neural network (exhaustive pruning)	73.45%
[24]	Financial statement fraud from Accounting and Auditing Enforcement Releases by the Securities and Exchange Commission	Genetic algorithm	98%-100%
[52]	Transactional fraud in automated bank machines and point of sale from a financial institution	Coevolution artificial immune system	95.862-97.122%
		Standard evolution artificial immune system	99.311%

Table 6. Qualitative analysis of methods researched in existing fraud detection literature

Method Investigated	Fraud Investigated	Research
Neural network	Financial statement fraud	[10], [28], [38]
Logistic model	Credit card fraud	[7]

	Insurance fraud	[36], [49], [6]
	Financial statement fraud	[38], [25]
Support vector machine	Credit card fraud	[7], [50]
	Financial statement fraud	[38], [25]
Decision trees, forests and CART	Credit card fraud	[7], [50], [40]
	Financial statement fraud	[10], [3], [28]
Genetic algorithm/programming	Financial statement fraud	[38], [24]
Text mining	Financial statement fraud	[20], [13], [56]
Group method of data handling	Financial statement fraud	[38]
Response-surface methodology	Financial statement fraud	[57]
Self-organizing map	Credit card fraud	[37], [41], [33]
Bayesian belief network	Corporate fraud	[23]
	Financial statement fraud	[28]
Process mining	Securities and commodities fraud	[27]
Artificial immune system	Credit card fraud	[52], [51], [43]
Hybrid methods	Credit card fraud	[35], [16]
	Insurance fraud	[6]
	Financial statement fraud	[13], [15], [26]
All/generic	All/generic	[55], [32]

Table 7. Methods studied by type of fraud investigated

Fraud Type	Method Applied	Research on the Type of Fraud
Credit card	Support vector machines Decision trees Self-organising maps Fuzzy logic Artificial immune system Hybrid methods	[7] and [34] investigated credit card fraud from an international operation [37] investigated a banking database from the Singapore branch of a well-known international bank [41] investigated fraud in multinational department stores [16] investigated typical consumer spending to determine fraud in a major bank in Turkey [35] investigated variation in legitimate customer transaction behaviour with synthesised credit card data [52] investigated automated bank machines and point of sale from an anonymous financial institution [50] investigated credit card transactions from two separate banks [51] investigated transactions from a major Australian bank [40] investigated six months of transactions from an anonymous

		bank [33] investigated credit card accounts of residents in Warsaw [43] investigated credit card transactions from a Brazilian bank.
Securities and commodities and other Corporate	Bayesian belief network Process mining	[27] investigated internal transactional fraud from a successful, anonymous European financial institution [23] Investigated emails and discussion group messages to detect corporate fraud
Insurance Fraud	Logistic model Hybrid methods	[36], [49] and [6] all investigated motor insurance claims from Spanish insurance companies
Financial statement	Response-surface methodology Neural networks Decision trees Bayesian belief networks Support vector machine Genetic algorithms Group method of data handling Logistic model (regression) Text mining Hybrid methods	[57] investigated financial statement fraud in general [28] investigated a selection of Greek manufacturing firms [38], [10], and [3] investigated a series of public Chinese companies [20] and [26] investigated managerial statements from official company documents [24] and [13] investigated Accounting and Auditing Enforcement Releases authored by a selection of US companies, and [15] investigated the same documents with a focus on Chinese companies. [56] investigated the litigation section of the Securities Exchange Commission website. [25] investigated Taiwanese companies that had been accused of fraud

Table 8. Most successful methods for each fraud type based on the accuracy measure

Fraud Investigated	Best Performer		Comparative Performers	
	Method	Accuracy	Method	Accuracy
Credit card fraud	Self-organising map	100%	Logistic regression	99.4%
			Support vector machine	99.6%
			Random forest	99.6%

			Artificial immune system	96.4%
Insurance fraud	Hybrid method	99.5%	Logistic regression	60.7%
Financial statement fraud	Neural network	98.1%	Decision tree	73.6%
			Bayesian belief network	90.3%
			Support vector machine	92.0%
			Genetic programming	94.1%
			Group method of data handling	93.0%
			Logistic regression	79.0%
			Hybrid method	95.7%
			Text mining	75.4%
			CDA	71.4%

Table 9. Datasets used in fraud detection research

Fraud Type	Research	Dataset
Credit card	[7], [34]	Almost 50 million credit card transactions from an international company from 2006 to 2007, all of which occurred in a single country.
	[37]	Approximately 100 transactions per customer for a single month at the Singaporean branch of an anonymous, well-known bank.
	[41]	Credit card transactions provided by a selection of retail companies in Chile, of which approximately 0.3% were fraudulent.
	[16]	Approximately 250000 credit card transactions from an anonymous Turkish bank, including 1050 that are fraudulent.
	[35]	Unknown amount of synthesised credit card transactional data.
	[52]	Selection of 522728 legitimate and 346 fraudulent credit card transactions from an anonymous financial institution.
	[50]	Credit card transactions from two banks during 2005, including 176 million transactions from nearly 17 million accounts of which 14281 experienced fraud.
	[51]	Total of 640361 transactions for 21746 credit cards from a major Australian bank
	[40]	Deliberate undersampling of approximately 11344000 credit card transactions over a six month period, 484 of which were legitimate.
	[33]	10000 credit card accounts from a bank in Warsaw from January to March 2005, of which 100 were fraudulent.

	[43]	Anonymous transaction records from a Brazilian bank, of which 3.74% were fraudulent in some way.
Securities and commodities and other Corporate	[27]	Random selection of 10000 process instances from an anonymous European financial institution in 2007.
	[23]	Random sample of 50 disgruntled and 40 non-disgruntled message obtained from electronic, intra-company communications.
Insurance Fraud	[36]	Selection of 4970 audited motor insurance claims from 2000, including 681 cases of fraud.
	[49]	Selection of. 2403 audited motor insurance claims from 2000, including 174 cases of fraud.
	[6]	Selection of 10000 audited motor insurance claims from 2000, including 101 cases of fraud.
Financial statement	[28]	Financial statements from 76 Greek manufacturing firms, of which 38 had demonstrated fraud.
	[38]	Financial statements from 202 companies listed on Chinese stock exchanges, of which 101 had demonstrated fraud.
	[10]	Financial statements from a selection of 202 Chinese companies.
	[3]	10 known Chinese companies that have been found guilty of fraud between 1999 and 2002
	[20]	10-K documents from 2006-2008 for various known-fraudulent US companies, matched with a selection of similar, legitimate companies.
	[26]	202 10-K documents for US companies from 1995-2004, of which half were fraudulent.
	[24]	Observations for 390 companies from US Accounting and Auditing Enforcement Releases between 1991 and 2004, of which 51 companies demonstrated fraud.
	[13]	10-K documents from 61 fraudulent and 61 non-fraudulent US companies from between 1993 and 2002.
	[56]	62 litigation releases from the Securities and Exchange Commission website.
	[25]	A series of companies prosecuted for fraud in Taiwan between 1998 and 2009, matched with similar legitimate companies in a ratio of 1:4
	[15]	10-K documents from 17 fraudulent and 17 non fraudulent Chinese companies from 2003 to 2014