Fraud Detection using Data Mining Techniques

Shivakumar Swamy N
Ph.D Scholar, Dept. of CSE
JJTU,Jhunjhunu,Rajastan-333001

Prof. Sanjeev C. Lingareddy
Prof. and Head, Dept. of CSE
Alpha College of Engineering, Bangalore

Abstract - Data mining technology is applied to fraud detection to establish the fraud detection model, describe the process of creating the fraud detection model, then establish data model with ID3 decision tree, and establish example of fraud detection model by using this model. As e-commerce sales continue to grow, the associated online fraud remains an attractive source of revenue for fraudsters. These fraudulent activities impose a considerable financial loss to merchants, making online fraud detection a necessity. The problem of fraud detection is concerned with not only capturing the fraudulent activities, but also capturing them as quickly as possible. This timeliness is crucial to decrease financial losses.

1.INTRODUCTION

Data mining is about finding insights which are statistically reliable, unknown previously, and actionable from data (Elkan, 2001). This data must be available, relevant, adequate, and clean.

Also, the data mining problem must be well-defined, cannot be solved by query and reporting tools, and guided by a data mining process model .

The term fraud here refers to the abuse of a profit organization's system without necessarily leading to direct legal consequences. In a competitive environment, fraud can become a business critical problem if it is very prevalent and if the prevention procedures are not fail-safe. Fraud detection, being part of the overall fraud control, automates and helps reduce the manual parts of a screening/checking process. This area has become one of the most established industry/government data mining applications.

Given the reality, the best cost effective option is to tease out possible evidences of fraud from the available data using mathematical algorithm. Evolved from numerous research communities, especially those from developed countries, the analytical engine within these solutions and software are driven by artificial immune systems, artificial intelligence, auditing, database, distributed and parallel computing, econometrics, expert systems, fuzzy logic, genetic algorithms, machine learning, neural networks, pattern recognition, statistics, visualization and others. There are plenty of specialized fraud detection solutions and software1 which protect businesses such as credit card, e-commerce, insurance, retail, telecommunications industries.

There are often two main criticisms of data mining-based fraud detection research: the dearth of publicly available real data to perform experiments on; and the lack of published well researched methods and techniques. To counter both of them, this paper garners all related literature for categorization and Comparison, selects some innovative methods and techniques for discussion; and points toward other data sources as possible alternatives.

 The primary objective of this paper is to define existing challenges in this domain for the different types of large data sets and streams. It categorizes, compares, and summarizes relevant data mining-based fraud detection methods and techniques in published academic and industrial research.

 The second objective is to highlight promising new directions from related adversarial data mining fields/applications such as epidemic/outbreak detection, insider trading, intrusion detection, money laundering, spam detection, and terrorist detection. Knowledge and experience from these adversarial domains can be interchangeable and will help prevent repetitions of common mistakes and "reinventions of the wheel".

Section 2 – Who are the white collar criminals which a fraud detection system should be designed to discover? Where can one apply data mining techniques to commercial fraud ?

Section 3 – What data is available for fraud detection? Which performance measurements are appropriate for analysis ?

Section 4 – Which techniques often used for automated fraud detection ? What combinations of techniques have been recommended? What are their weaknesses ?

Section 5 – What analytical methods and techniques from other adversarial domains can one apply in fraud detection ?

Section 6 – How is this fraud detection survey different from others ?

Section 7 - Concludes with a brief summary. 1School of Business Systems, Faculty of Information Technology, Monash University, Clayton campus, Wellington Road, Clayton, Victoria 3800, Australia

II. BACKGROUND

This section highlights the types of fraudsters.

2.1 Fraudsters

Traditionally, each business is always susceptible to internal fraud or corruption from its management (high-level) and non-management employees (low-level). In addition to internal and external audits for fraud control, data mining can also be utilized as an analytical tool. the fraudster can be an external party, or parties. Also, the fraudster can either commit fraud in the form of a prospective/existing customer (consumer) or a prospective/existing supplier (provider). The external fraudster has three the average offender, criminal offender, and organized crime offender. Average offenders display random and/or occasional dishonest behavior when there is opportunity, sudden temptation, or when suffering from financial hardship.

In contrast, the more risky external fraudsters are individual criminal offenders and organized/group crime offenders (professional/career fraudsters) because they repeatedly disguise their true identities and/or evolve their *modus operandi* over time to approximate legal forms and to counter detection systems.

Therefore, it is important to account for the strategic interaction, or moves and countermoves, between a fraud detection system's algorithms and the professional fraudsters' *modus operandi*.

It is probable that internal and insurance fraud is more likely to be committed by average offenders; credit and telecommunications fraud is more vulnerable to professional fraudsters.

III. DATA AND MEASUREMENTS

This section discusses the types of available data and previously used performance measures.

3.1 Structured data

This subsection aims to define the attributes and examples which have been used for previous fraud detection experimental studies and actual systems. By doing so, future studies on fraud detection will find this useful to either validate their real data or create synthetic data.

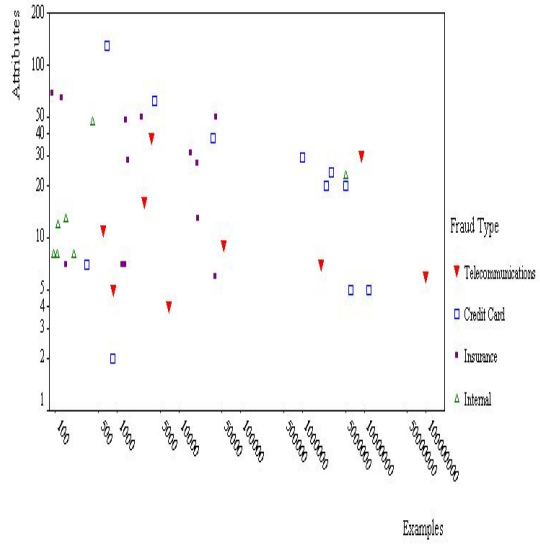


Figure 3.1: Scatter plot of the data

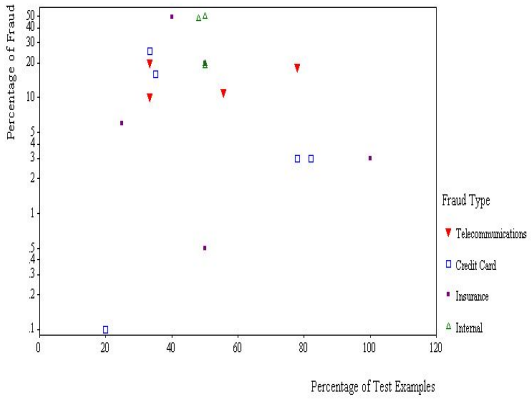


Figure 3.2: Scatter plot of % of fraud and % of test

Fig 3.1 shows the number of original attributes (vertical axis) and pre-sampled examples (horizontal axis) from internal, insurance, credit card, and telecommunications fraud detection literature. Generally, attributes can be binary, numerical (interval or ratio Scales), categorical (nominal or ordinal scales), or a mixture of the three. 16 data sets have less than 10 attributes, 18 data sets have between 10 to 49 attributes, 5 data sets have between 50 to 99 attributes, and only 1 data set used more than 100 attributes (Wheeler and Aitken, 2000).

Management data sets are the smallest (all have less than 500 examples), except for employee/retail data with more than 5 million transactions (Kim *et al*, 2003). Insurance data sets consist of hundreds of examples and the largest contain 40000 examples (Williams, 1999). Most credit transactional data have more than 1 million transactions and the largest contain more than 12 million transactions per year (Dorronsoro *et al*, 1997). Telecommunications data are the largest because they comprise of transactions generated by hundreds,

thousands, or millions of accounts. The largest reported is produced by at least 100 million telecommunications accounts (Cortes *et al*, 2003)

Fig 3.2 shows the % of fraud (vertical axis) and % of test examples (horizontal axis) of the entire data set described in each study. Six studies using credit transactional and insurance data have less than 10% fraud. In particular, Foster and Stine (2004) and Bentley (2000) have as low as 0.1 % fraud in credit transactional data and 0.5 % fraud in home insurance data respectively. More than 80% (16 papers) of the 19 papers has skewed data with less than 30% fraud. The average of the proportion of test examples to total examples of the 19 papers is around 50%. The specific attributes used for detecting each fraud type are generally the same

3.2 Performance Measures

Most fraud departments place monetary value on predictions to maximise cost savings/profit and according to their policies. They can either define explicit cost (Phua *et al*, 2004; Chan *et al*, 1999) or benefit models (Fan *et al*, 2004; Wang *et al*, 2003;Cahill *et al*, 2002) suggests giving a score for an instance (phone call) by determining the similarity of it to known fraud examples (fraud styles) divided by the dissimilarity of it to known legal examples (legitimate telecommunications account). Most of the fraud detection studies using supervised algorithms since 2001 have abandoned measurements such as true positive rate (correctly detected fraud divided by actual fraud) and accuracy at a chosen threshold (number of instances predicted correctly, divided by the total number of instances).

In fraud detection, misclassification costs (false positive and false negative error costs) are unequal, uncertain, can differ from example to example, and can change over time. In fraud detection, a false negative error is usually more costly than a false positive error. Regrettably, some recent studies on credit card transactional fraud (Chen *et al*, 2004) and telecommunications superimposed fraud(Kim *et al*, 2003) still aim to only maximize accuracy. Some use Receiver Operating Characteristic (ROC) analysis (true positive rate versus false positive rate).

Apart from Viaene *et al* (2004), no other fraud detection study on supervised algorithms has sought to maximise Area under the Receiver Operating Curve (AUC) and minimise cross entropy (CXE). AUC measures how many times the instances have to be swapped with their neighbours when sorting data by predicted scores; and CXE measures how close predicted scores are to target scores. In addition, Viaene *et al* (2004) and Foster and Stine (2004) seek to minimise Brier score (mean squared error of predictions). Caruana and Niculescu-Mizil (2004) argues that the most effective way to assess supervised algorithms is to use one metric from threshold, ordering, and probability metrics; and they justify using the average of mean squared error, accuracy, and AUC. Fawcett and Provost (1999) recommend Activity Monitoring Operating Characteristic (AMOC) (average score versus false alarm rate) suited for timely credit transactional and telecommunications superimposition fraud detection.

For semi-supervised approaches such as anomaly detection, Lee and Xiang (2001) propose entropy, conditional entropy, relative conditional entropy, information gain, and information cost. For unsupervised algorithms, Yamanishi *et al* (2004) used the Hellinger and logarithmic scores to find statistical outliers for insurance; Burge and Shawe-Taylor (2001) employed Hellinger score to determine the difference between short-term and long term profiles for the telecommunications account. Bolton and Hand (2001) recommends the *t*-statistic as a score to compute the standardized distance of the target account with centroid of the peer group; and also to detect large spending changes within accounts.

IV. METHODS AND TECHNIQUES

This section examines four major methods commonly used, and their corresponding techniques and algorithms.

4.1 Overview

Figure 4.1 shows that many existing fraud detection systems typically operate by adding fraudulent/claims/applications/transactions/accounts/sequenc(A) to “black lists” to match for likely frauds in the new instances(E). Some use hard-coded rules which each transaction should meet such as matching addresses and phone numbers, and price and amount limits (Sherman, 2002).

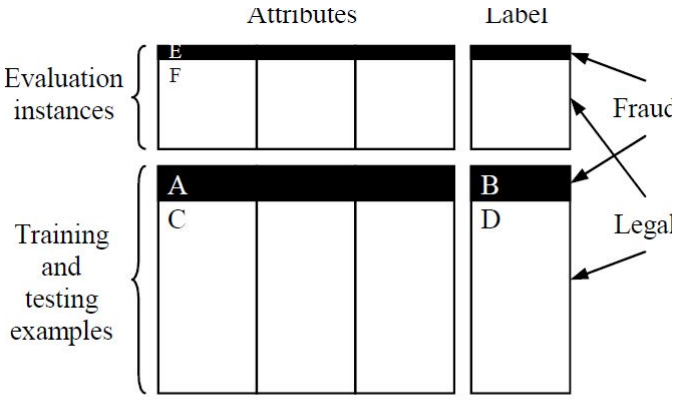


Figure 4.1: Structured diagram of the possible data for analysis. Dat
ning approaches can utilise training/testing data with labels, only le
examples, and no labels to predict/describe the evaluation data.

With reference to Figure 4.1, the common data mining approaches to determine the most suspicious examples from the incoming data stream (evaluation data) are:

1. Supervised approaches
2. Unsupervised approaches
3. Hybrid approaches
4. Semi-supervised approaches

Labelled training data (A + B + C + D) can be processed by single *supervised* algorithms (Section 4.2). A better suggestion is to employ hybrids such as multiple supervised algorithms (Section 4.3.1), or both supervised and unsupervised algorithms (Section 4.3.2) to output suspicion scores, rules and/or visual anomalies on evaluation data.

Therefore it is necessary to Combine training data (the class labels are not required here) with evaluation data (A + C + E + F). These should be processed by single or multiple *unsupervised* algorithms to output suspicion scores, rules and/or visual anomalies on evaluation data.

4.2 Supervised Approaches on Labelled Data (A + B + C + D)

Predictive supervised algorithms examine all previous labeled transactions to mathematically determine how a standard fraudulent transaction looks like by assigning a risk score(Sherman, 2002). Neural networks are popular and support vector machines (SVMs) have been applied. Barse *et al* (2003) used a multi-layer neural network with exponential trace memory to handle temporal dependencies in synthetic Video-on-Demand log data. Syeda *et al* (2002) propose fuzzy neural networks on parallel machines to speed up rule production for customer-specific credit card fraud detection. Kim *et al* (2003) proposes SVM ensembles with either bagging and boosting with aggregation methods for telecommunications subscription fraud.

The neural network and Bayesian network comparison study (Maes *et al*, 2002) uses the STAGE algorithm for Bayesian networks and back propagation algorithm for neural networks in credit transactional fraud detection. Comparative results show that Bayesian networks were more accurate and much faster to train, but Bayesian networks are slower when applied to new instances.

Other techniques include expert systems, association rules, and genetic programming. Expert systems have been applied to insurance fraud. Major and Riedinger (2002) have implemented an actual five-layer expert system in which expert knowledge is integrated with statistical information assessment to identify medical insurance The above supervised algorithms are conventional learning techniques which can only process structured data from single 1- to-1 data tables. Further research using labelled data in fraud detection can benefit from applying relational learning approaches such as Inductive Logic Programming (ILP).

4.3 Hybrid Approaches with Labelled Data

4.3.1 Supervised Hybrids (A + B + C + D)

Popular supervised algorithms such as neural networks, Bayesian networks, and decision trees have been combined or applied in a sequential fashion to improve results. Phua *et al*(2004) proposes back propagation neural networks, naive Bayes, and C4.5 as base classifiers on data partitions derived from minority oversampling with replacement. Its originality lies in the use of a single meta-classifier (stacking) to choose the best base classifiers, and then combine these base classifiers' predictions (bagging) to produce the best cost savings on automobile insurance claims. Ormerod *et al*(2003) recommends a rule generator to refine the weights of the Bayesian network. Kim and Kim(2002) propose a decision tree to partition the input space, tanh as a

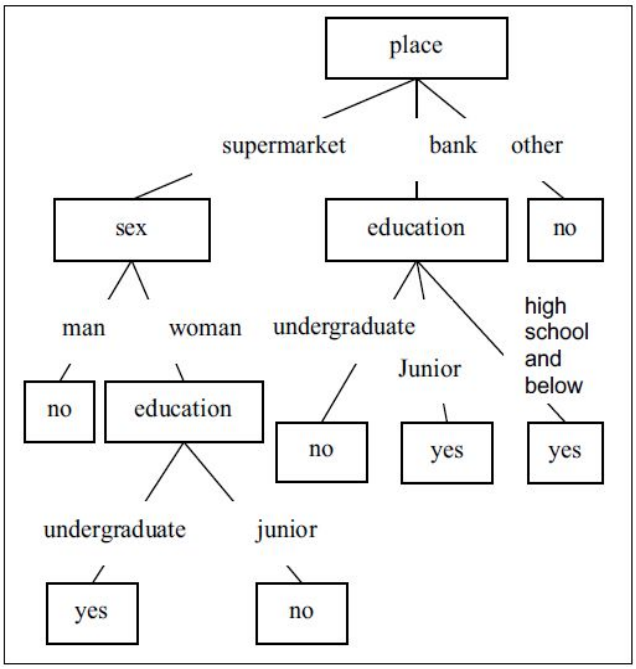
weighting function to generate fraud density, and subsequently a backpropagation neural network to generate a weighted suspicion score on credit card transactions.

Also, He *et al*(1999) propose genetic algorithms to determine optimal weights of the attributes, followed by *k*-nearest neighbor algorithm to classify the general practitioner data. They claim significantly better results than without feature weights and when compared to CBR.

4.3.2 Supervised/Unsupervised Hybrids (A + B + C+ D)

There is extensive work on labelled data using both supervised and unsupervised algorithms in telecommunications fraud detection. Cortes and Pregibon (2001) propose the use of signatures (telecommunication account summaries) which are updated daily (time-driven). Fraudulent signatures are added to the training set and processed by supervised algorithms such as a tree, slipper, and model-averaged regression. The authors remark that fraudulent toll-free numbers tend to have extensive late night Activity and long call durations. Cortes and Pregibon (2001) use signatures assumed to be legitimate to detect significant changes in calling behaviour. Association rules is used to discover interesting country combinations and temporal information from the previous month. A graph-theoretic method (Cortes *et al*, 2003) is used to visually detect communities of interest of fraudulent international call accounts (see Section 4.5). Cahill *et al* (2002) assign an averaged suspicion score to each call (event-driven) based on its similarity to fraudulent signatures and dissimilarity to its account's normal signature. Calls with low scores are used to update the signature and recent calls are weighted more heavily than earlier ones in the signature.

Fawcett et al(1997) present fraud rule generation from each cloned phone account's labelled data and rule selection to cover most accounts. Each selected fraud rule is applied in the form of monitors (number and duration of calls) to the daily legitimate usage of each account to find anomalies. The selected monitors' output and labels on an account's previous daily behaviour are used as training data for a simple Linear Threshold Unit. An alarm will be raised on that account if the suspicion score on the next evaluation day exceeds its threshold. In terms of cost savings and accuracy, this method performed better than other methods such as expert systems, classifiers trained without account context, high usage, collision detection, velocity checking, and dialled digit analysis on detecting telecommunications superimposed fraud.



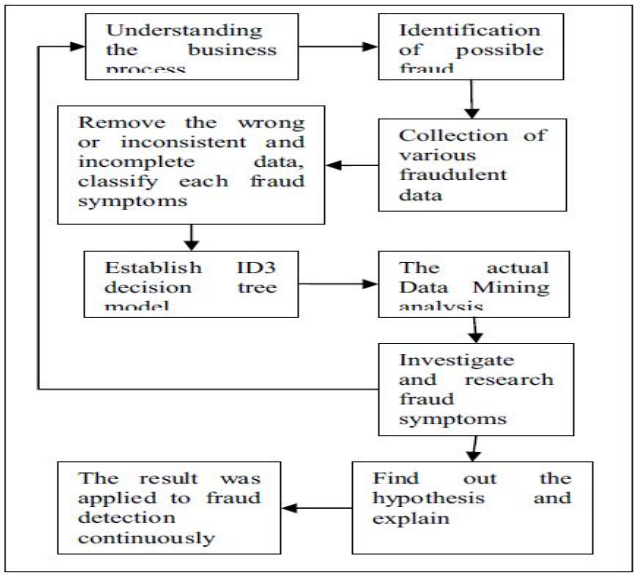
Decision tree

Two studies on telecommunications data show that supervised approaches achieve better results than unsupervised ones. With AUC as the performance measure, Moreau *et al* (1999) show that supervised neural network and rule induction algorithms outperform two forms of unsupervised neural networks which identify differences between short-term and long-term statistical account behavior profiles. The best results are from a hybrid model which combines these four techniques using logistic regression. Using true positive rate with no false positives as the performance measure, Taniguchi *et al* (1998) claim that supervised neural networks and Bayesian networks on labeled achieve significantly better outcomes than unsupervised techniques such as

Gaussian mixture models on each non-fraud user to detect anomalous phone calls. Unsupervised approaches have been used to segment the insurance data into clusters for supervised approaches. Williams and Huang (1997) applies a three step process: *k*-means for cluster
detection, C4.5 for decision tree rule induction, and domain knowledge, statistical summaries and visualisation tools for rule evaluation. Williams (1999) use a genetic algorithm, instead of C4.5, to generate rules and to allow the domain user, such as a fraud specialist, to explore the rules and to allow them to evolve accordingly on medical insurance claims.

4.4 Semi-supervised Approaches with Only Legal (Non-fraud) Data (C)

Kim *et al* (2003) implements a novel fraud detection method in five steps: First, generate rules randomly using association rules algorithm *Apriori* and increase diversity by a calendar schema; second, apply rules on known legitimate action database, discard any rule which matches this data; third, use Murad and Pinkas(1999) use profiling at call, daily and overall Levels of normal behavior from each telecommunications account. The common daily profiles are extracted using a clustering algorithm with cumulative distribution distance function. An alert is raised if the daily profile's call duration, destination, and quantity exceed the threshold and standard deviation of the overall profile. Aleskerov *et al* (1997) experiment with auto-associative neural networks (one hidden layer and the same number of input and output neurons) on each credit card account's legal transactions. Kokkinaki (1997) proposes similarity trees (decision trees with Boolean logic functions) to profile each legitimate customer's behavior to detect deviations from the norm and cluster analysis to segregate each legitimate customer's credit card transactions.



FD MODEL

4.5 Unsupervised Approaches with Unlabelled Data (A + C + E + F)

Link analysis and graph mining are hot research topics in antiterrorism, law enforcement, and other security areas, but these techniques seem to be relatively under-rated in fraud detection research. A white paper (NetMap, 2004) describes how the emergent group algorithm is used to form groups of tightly connected data and how it led to the capture of an actual elusive fraudster by visually analysing twelve months worth of insurance claims. There is a brief application description of a visual telecommunications fraud detection system (Cox, 1997) which flexibly encodes data using colour, position, size and other visual characteristics with multiple different views and levels. The intuition is to combine human detection with machine computation. Cortes *et al* (2001) examines temporal evolution of large dynamic graphs' for telecommunications fraud detection. Each graph is made up of sub graphs called Communities Of Interest (COI). To overcome instability of using just the current graph, and storage and weightage problems of using all graphs at all time steps; the authors used the exponential weighted average approach to update sub graphs daily. By linking mobile phone accounts using call quantity and durations to form COIs, the authors confirm two distinctive characteristics of fraudsters. First, fraudulent phone accounts are linked - fraudsters call each other or the same phone numbers. Second, fraudulent call behavior from flagged frauds are reflected in some new phone accounts - fraudsters retaliate with application fraud/identity crime after being detected. Cortes *et al* (2003) states their contribution to dynamic graph research in the areas of scale, speed, dynamic updating, condensed representation of the graph, and measure direct interaction between nodes.

Some forms of unsupervised neural networks have been applied. Dorronsoro *et al* (1997) creates a non-linear discriminant analysis algorithm which do not need labels. It minimizes the ratio of the determinants of the within and between class variances of weight projections. There is no history on each credit card account's past

transactions, so all transactions have to be segregated into different geographical locations. The authors explained that the installed detection system has low false positive rates, high cost savings, and high computational efficiency. Burge and Shawe- Taylor (2001) use a recurrent neural network to form short-term and long-term statistical account behaviour profiles. Hellinger distance is used to compare the two probability distributions and give a suspicion score on telecommunications toll tickets.

In addition to cluster analysis (Section 4.3.2), unsupervised approaches such as outlier detection, spike detection, and other forms of scoring have been applied. Yamanishi *et al* (2004) demonstrated the unsupervised SmartSifter algorithm which can handle both categorical and continuous variables, and detect statistical outliers using Hellinger distance, on medical insurance data. Bolton and Hand (2001) recommend Peer Group Analysis to monitor inter-account behavior over time. It compares the cumulative mean weekly amount between a target account and other similar accounts (peer group) at subsequent time points. The distance metric/suspicion score is a *t*-statistic which determines the standardised distance from the centroid of the peer group. The time window to calculate peer group is thirteen weeks and future time window is four weeks on credit card accounts. Bolton and Hand (2001) also suggest Break Point Analysis to monitor intra account behavior over time. It detects rapid spending or sharp increases in weekly spending within a single account. Accounts are ranked by the *t*-test. The fixed-length moving transaction window contains twenty-four transactions: first twenty for training and next four for evaluation on credit card accounts. Brockett *et al* (2002) recommends Principal Component Analysis of RIDIT scores for rank-ordered categorical attributes on automobile insurance data. Hollmen and Tresp (1998) present an experimental real-time fraud detection system based on a Hidden Markov Model (HMM).

4.6 Critique of Methods and Techniques

 In most scenarios of real-world fraud detection, the choice of data mining techniques is more dependent on the practical issues of operational requirements, resource constraints, and management commitment towards reduction of fraud than the technical issues poised by the data.

 Other novel commercial fraud detection techniques include graph-theoretic anomaly detection2 and Inductive Logic Programming3. There has not been any empirical evaluation of commercial data mining tools for fraud detection since Abbott *et al* (1998).

 There is too much emphasis by research on complex, nonlinear supervised algorithms such as neural networks and support vector machines. In the long term, less complex and faster algorithms such as naive Bayes (Viaene *et al*, 2002) and logistic regression (Lim *et al*, 2000) will produce equal, if not better results (see Section 3.2), on population-3drifting, concept-drifting, adversarial-ridden data. If the incoming data stream has to be processed immediately in an event-driven system or labels are not readily available, then semisupervised and unsupervised approaches are the only data mining options.

 Other related data mining techniques covered by survey papers and bibliographies include outlier detection (Hodge and Austin, 2004), skewed/imbalanced/rare classes4 (Weiss, 2004), sampling (Domingos *et al*, 2002), cost sensitive learning5, stream mining6, graph mining (Washio and Motoda, 2003), and scalability (Provost and Kolluri, 1999)

.

V. OTHER ADVERSARIAL DOMAINS

This section explains the relationship between fraud detection. Three other similar domains.

5.1 Terrorist Detection

There had been simplistic technical critiques of data mining for terrorist detection such as low accuracy (unacceptably high false positive rates in skewed data) and serious privacy violations (massive information requirements). To counter them, Jensen *et al* (2003) recommend fixed-size clustering to generate true class labels and the linked structure of data. Scores are randomly drawn from either the negative or positive entities' normal distributions.

The second-round classifier averages an entity's first-round score and scores of all its neighbours. To reduce false positives, results show that second-round classifier reduces false positive rates while maintaining true positive rates of first-round classifier. To reduce information requirements, results show moderately high accuracy through the use of only twenty percent of the data. Surveillance systems for terrorist, bio-terrorist, and chemo terrorist detection often depend on spatial and spatio temporal data. These are unsupervised techniques highly applicable to fraud detection. Neill and Moore (2004) employ Kulldorff's spatial scan statistic and the overlap-kd tree data structure. It efficiently finds the most significant densities from latitude and longitude of patient's home in real emergency department, and zip codes in retail cough and cold medication sales data. Das *et al* (2004) utilise Partially Observable Markov Decision Process (POMDP) with Kulldorff's spatial scan statistic on to detect artificial attacks from real emergency department's spatio-temporal data.

Bio-terrorism detection aims to detect irregularities in temporal data. Similar to fraud detection, data has to be partially simulated by injecting epidemics, and performance is evaluated with detection time and number of false positives. Wong *et al* (2003) apply Bayesian networks to uncover simulated anthrax attacks from real emergency department data. Hutwagner *et al* (2003) describe the use of cumulative sum of deviations

in the Early Aberration Reporting System (EARS). Goldenberg *et al* (2002) use time series analysis to track early symptoms of synthetic anthrax outbreaks from daily sales of retail medication (throat, cough, and nasal) and some grocery items (facial tissues, orange juice, and soup).

5.2 Financial Crime Detection

Financial crime here refers to money laundering, violative trading, and insider trading and the following are brief application descriptions which correspond to each type of monitoring system temporal relationships between events from market data which exists in a potential violation pattern. Association rules and decision trees are used to discover new patterns or refined rules which reflect
behavioural changes in the marketplace. It has been successfully

It mines for explicit and implicit relationships among the entities and events, all of which form episodes or scenarios with specific identifiers. It has been reported to be successful in generating breaks the main stock markets for insider trading (trading upon inside information of a material nature) and misrepresentation fraud (falsified news).

Use of large amounts of unstructured text and web data such as on Correlation Analysis (LDCA) which uses a correlation measure with fuzzy logic to determine similarity of patterns between thousands of paired textual items which have no explicit links. It comprises of link hypothesis, link generation, and link identification based on financial transaction timeline analysis to generate community models for the prosecution of money laundering criminals.

.

5.3 Intrusion and Spam Detection

There are multiple data sources for intrusion detection and the common ones are at host level, network level, and user level. Otey the benchmark KDD cup 1999 network intrusion detection data is often used. In addition, semi-real user level data are common, the “intrusions” are usually simulated using another user data and 10 “real” part refers to normal computer usage data for each legitimate user.

In intrusion detection terms, misuse detection is for matching known attacks (using A of Figure 4.1); and anomaly detection is for discovering unknown attacks (using C of Figure 4.1 and see Section 4.4). The current research in both intrusion detection and spam detection are on anomaly detection (semi-supervised) and unsupervised approaches. In intrusion detection research, the use of clustering to reduce data and HMMs for anomaly detection had been popular. Lane and Brodley (2003) detail that *k*-means to compress data and report that HMMs performed slightly better than instance-based learning (IBL) for semi-real user level data.

Similarly, Cho (1999) use SOM to decrease data for HMMmodelling. The author show that multiple HMM models with fuzzy logic can be used to reduce false positive rates. Also, Stolfo the authors comment that SVM is the best supervised algorithm but the detection time is too long for an event-driven system.

The use of game theory to model the strategic interaction between the system and adversary has been recently introduced into intrusion and spam detection research. Patcha and Park (2004). Tested under different false positives costs, the game-theoretic naive Bayes classifier outperforms the conventional classifier by efficiently predicting no false positives with relatively low false negatives.

VI. RELATED WORK

This paper examines fraud detection from a practical data oriented, performance-driven perspective rather than the typical application-oriented or technique-oriented view of the three other recent survey papers. In addition, this survey clearly defines the underlying technical problems and covers more relevant fraud types, methods, and techniques than any of the other survey papers. For example, internal fraud and the various hybrid approaches are presented here. Also, some criticisms of the current fraud detection field are given and possible future contributions to data mining-based fraud detection from related domains are highlighted.

VII. CONCLUSION & FUTURE WORK

This survey has explored almost all published fraud detection. It defines the adversary, the types and subtypes of fraud, the technical nature of data, performance metrics, and the methods and techniques. After identifying the limitations in methods and techniques of fraud detection, this paper shows that this field can benefit from other related fields. Specifically, unsupervised approaches from counter terrorism work, actual monitoring systems and text mining from law enforcement, and semi supervised and game-theoretic approaches from intrusion and spam detection communities can contribute to future fraud detection research. However, Fawcett and Provost (1999) show that there are no guarantees when they successfully applied their fraud detection method to news story monitoring but unsuccessfully to intrusion detection. Future work will be in the form of credit application fraud detection.

REFERENCES

[1] Abbott, D., Matkovsky, P. & Elder, J. (1998). An Evaluation of High-End Data Mining Tools for Fraud Detection. Proc. of IEEE SMC98.
[2] Aleskerov, E., Freisleben, B. & Rao, B. (1997). CARDWATCH: A Neural Network-Based Database Mining System for Credit Card Fraud Detection. Proc. of the IEEE/IAFE on Computational Intelligence for Financial Engineering, 220-226.
[3] Artis, M., Ayuso M. & Guillen M. (1999). Modelling Different Types of Automobile Insurance Fraud Behaviour in the Spanish Market. Insurance Mathematics and Economics 24: 67-81.
[4] Barse, E., Kvarnstrom, H. & Jonsson, E. (2003). Synthesizing Test Data for Fraud Detection Systems. Proc. of the 19th Annual Computer Security Applications Conference, 384-395.
[5] 5 .Belhadji, E., Dionne, G. & Tarkhani, F. (2000). A Model for the Detection of Insurance Fraud. The Geneva Papers on Risk and
[6] Insurance 25(4): 517-538.
[7] Bell, T. & Carcello, J. (2000). A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting. Auditing: A Journal of Practice and Theory 10(1): 271-309.
[8] Bentley, P. (2000). Evolutionary, my dear Watson: Investigating Committee-based Evolution of Fuzzy Rules for the Detection of Suspicious Insurance Claims. Proc. Of GECCO2000.
[9] Bentley, P., Kim, J., Jung., G. & Choi, J. (2000). Fuzzy Darwinian Detection of Credit Card Fraud. Proc. of 14th Annual Fall Symposium Of the Korean Information Processing Society.
[10] Bolton, R. & Hand, D. (2002). Statistical Fraud Detection: A Review (With Discussion). Statistical Science 17(3): 235-255.
[11] Bolton, R. & Hand, D. (2001). Unsupervised Profiling Methods for Fraud Detection. Credit Scoring and Credit Control VII.
[12] Brockett, P., Derrig, R., Golden, L., Levine, A. & Alpert, M.(2002). Fraud Classification using Principal Component Analysis of RIDITs. Journal of Risk and Insurance 69(3): 341-371.
[13] Burge, P. & Shawe-Taylor, J. (2001). An Unsupervised Neural Network Approach to Profiling the Behaviour of Mobile Phone Users for Use in Fraud Detection. Journal of Parallel and Distributed Computing 61: 915-925.
[14] Cahill, M., Chen, F., Lambert, D., Pinheiro, J. & Sun, D (2002). Detecting Fraud in the Real World. Handbook of Massive Datasets 911-930.
[15] Caruana, R. & Niculescu-Mizil, A. (2004). Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria. Proc. of SIGKDD04, 69-78.
[16] Chan, P., Fan, W., Prodromidis, A. & Stolfo, S. (1999). Distributed Data Mining in Credit Card Fraud Detection. IEEE Intelligent Systems 14: 67-74.
[17] Chen, R., Chiu, M., Huang, Y. & Chen, L. (2004). Detecting Credit Card Fraud by Using Questionaire-Responded Transaction Model Based on Support Vector Machines. Proc. of IDEAL2004, 800-806.
[18] Chiu, C. & Tsai, C. (2004). A Web Services Based Collaborative Scheme for Credit Card Fraud Detection. Proc. of 2004 IEEE International Conference on e-Technology, e-Commerce and e- Service.
[19] Cortes, C., Pregibon, D. & Volinsky, C. (2003). Computational Methods for Dynamic Graphs. Journal of Computational and Graphical Statistics 12: 950-970.
[20] Cortes, C. & Pregibon, D. (2001). Signature-Based Methods for Data Streams. Data Mining and Knowledge Discovery 5: 167-182.
[21] Cox, E. (1995). A Fuzzy System for Detecting Anomalous Behaviors in Healthcare Provider Claims. In Goonatilake, S. & Treleaven, P.(eds.) Intelligent Systems for Finance and Business, 111-134. John Wiley and Sons Ltd.
[22] Das, K., Moore, A. & Schneider, J. (2004). Belief State Approaches to Signaling Alarms in Surveillance Systems. Proc. Of SIGKDD04, 539-544.
[23] Domingos, C., Gavalda, R. & Watanabe, O. (2002). Adaptive Sampling Methods for Scaling Up Knowledge Discovery Algorithms. Data Mining and Knowledge Discovery 6: 131-152.
[24] Dorronsoro, J., Ginel, F., Sanchez, C. & Cruz, C. (1997). Neural Fraud Detection in Credit Card Operations. IEEE Transactions On Neural Networks 8(4): 827-834.
[25] Elkan, C. (2001). Magical Thinking in Data Mining: Lessons from CoIL Challenge 2000. Proc. of SIGKDD01, 426-431.
[26] Ezawa, K. & Norton, S. (1996). Constructing Bayesian Networks to Predict Uncollectible Telecommunications Accounts. IEEE Expert October: 45-51.
[27] Fan, W. (2004). Systematic Data Selection to Mine Concept- Drifting Data Streams. Proc. of SIGKDD04, 128-137. 12
[28] Fanning, K., Cogger, K. & Srivastava, R. (1995). Detection of Management Fraud: A Neural Network Approach. Journal of Intelligent Systems in Accounting, Finance and Management 4:113-126.
[29] Fawcett, T. (1997). AI Approaches to Fraud Detection and Risk Management: Papers from the 1997 AAAI Workshop. Technical Report WS-97-07. AAAI Press.
[30] Fawcett, T. & Provost, F. (1999). Activity monitoring: Noticing Interesting Changes in Behavior. Proc. of SIGKDD99, 53-62.
[31] Fawcett, T. & Provost, F. (1997). Adaptive Fraud Detection. Data Mining and Knowledge Discovery 1(3): 291-316.
[32] Foster, D. & Stine, R. (2004). Variable Selection in Data Mining: Building a Predictive Model for Bankruptcy. Journal of American Statistical Association 99: 303-313.
[33] Goldberg, H., Kirkland, J., Lee, D., Shyr, P. & Thakker, D. (2003). The NASD Securities Observation, News Analysis & Regulation System (SONAR). Proc. of IAAI03.
[34] Goldenberg, A., Shmueli, G., Caruana, R. & Fienberg, S. (2002).Early Statistical Detection of Anthrax Outbreaks by Tracking Over-the-Counter Medication Sales. Proc. of the National Academy of Sciences, 5237-5249.
[35] Hawkins, S., He, H., Williams, G. & Baxter, R. (2002). Outlier Detection Using Replicator Neural Networks. Proc. Of DaWaK2002, 170-180.
[36] He, H., Graco, W. & Yao, X. (1999). Application of Genetic Algorithms and k-Nearest Neighbour Method in Medical Fraud Detection. Proc. of SEAL1998, 74-81.
[37] Hodge, V. & Austin, J. (2004). A Survey of Outlier Detection Methodologies. Artificial Intelligence Review 22: 85-126.
[38] Hollmen, J. & Tresp, V. (1998). Call-based Fraud detection in Mobile Communication Networks using a Hierarchical Regime-Switching Model. Proc. of Advances in Neural Information Processing Systems.
[39] Hutwagner, L., Thompson, W. & Seeman, M. (2003). The Bioterrorism Preparedness and Response Early Abberation Reporting System (EARS). Journal of Urban Health: Bulletin of the New York Academy of Medicine 80(2): 89-96.
[40] Jensen, D., Rattigan, M. & Blau, H. (2003). Information Awareness: A Prospective Technical Assessment. Proc. Of SIGKDD03, 378-387.
[41] Kim, H., Pang, S., Je, H., Kim, D. & Bang, S. (2003). Constructing Support Vector Machine Ensemble. Pattern Recognition 36: 2757-2767. 747-752.