Analysis on Credit Card Fraud Detection Methods

[1]S. Benson Edwin Raj, [2]A. Annie Portia
[1]Assistant Professor (SG), P.G., [2]Scholar
Department of CSE
Karunya University, Coimbatore

Abstract— **Due to the rise and rapid growth of E-Commerce, use of credit cards for online purchases has dramatically increased and it caused an explosion in the credit card fraud. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In real life, fraudulent transactions are scattered with genuine transactions and simple pattern matching techniques are not often sufficient to detect those frauds accurately. Implementation of efficient fraud detection systems has thus become imperative for all credit card issuing banks to minimize their losses. Many modern techniques based on Artificial Intelligence, Data mining, Fuzzy logic, Machine learning, Sequence Alignment, Genetic Programming etc., has evolved in detecting various credit card fraudulent transactions. A clear understanding on all these approaches will certainly lead to an efficient credit card fraud detection system. This paper presents a survey of various techniques used in credit card fraud detection mechanisms and evaluates each methodology based on certain design criteria.**

***Index Terms*—Electronic Commerce, Credit card fraud, Artificial Intelligence, Artificial Neural Networks, Sequence Alignment, Machine Learning.**

I. INTRODUCTION

The Credit Card Is A Small Plastic Card Issued To Users As A System Of Payment. It Allows Its Cardholder To Buy Goods And Services Based On The Cardholder's Promise To Pay For These Goods And Services. Credit Card Security Relies On The Physical Security Of The Plastic Card As Well As The Privacy Of The Credit Card Number. Globalization And Increased Use Of The Internet For Online Shopping Has Resulted In A Considerable Proliferation Of Credit Card Transactions Throughout The World. Thus A Rapid Growth In The Number Of Credit Card Transactions Has Led To A Substantial Rise In Fraudulent Activities. Credit Card Fraud Is A Wide-Ranging Term For Theft And Fraud Committed Using A Credit Card As A Fraudulent Source Of Funds In A Given Transaction. Credit Card Fraudsters Employ A Large Number Of Techniques To Commit Fraud. To Combat The Credit Card Fraud Effectively, It Is Important To First Understand The Mechanisms Of Identifying A Credit Card Fraud. Over The Years Credit Card Fraud Has Stabilized Much Due To Various Credit Card Fraud Detection And Prevention Mechanisms.

II. RELATED WORKS

Fraud detection involves monitoring the behavior of users in order to estimate, detect, or avoid undesirable behavior. To counter the credit card fraud effectively, it is necessary to understand the technologies involved in detecting credit card frauds and to identify various types of credit card frauds [20] [21] [22] . There are multiple algorithms for credit card fraud detection [21] [29]. They are artificial neural-network models which are based upon artificial intelligence and machine learning approach [5] [7] [9] [10] [16], distributed data mining systems [17] [19], sequence alignment algorithm which is based upon the spending profile of the cardholder [1] [6], intelligent decision engines which is based on artificial intelligence [23], Meta learning Agents and Fuzzy based systems [4]. The other technologies involved in credit card fraud detection are Web Services-Based Collaborative Scheme for Credit Card Fraud Detection in which participant banks can share the knowledge about fraud patterns in a heterogeneous and distributed environment to enhance their fraud detection capability and reduce financial loss [8] [13], Credit Card Fraud Detection with Artificial Immune System [13] [26], CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection [18] which is bases upon data mining approach [17] and neural network models, the Bayesian Belief Networks [25] which is based upon artificial intelligence and reasoning under uncertainty will counter frauds in credit cards and also used in intrusion detection [26], case-based reasoning for credit card fraud detection [29], Adaptive Fraud Detection which is based on Data Mining and Knowledge Discovery [27], Real-time credit card fraud using computational intelligence [28], and Credit card fraud detection using self-organizing maps [30]. Most of the credit card fraud detection systems mentioned above are based on artificial intelligence, Meta learning and pattern matching.

This paper compares and analyzes some of the good techniques that have been used in detecting credit card fraud. It focuses on credit card fraud detection methods like Fusion of Dempster Shafer and Bayesian learning [2][5][12][15][25], Hidden Markov Model [3], Artificial neural networks and Bayesian Learning approach [5][25],BLAST and SSAHA Hybridization[1][6][11][14][24], Fuzzy Darwinian System[4]. Section II gives an overview about those techniques. Section III presents a comparative survey of those techniques and section IV summarizes the fraud detection techniques.

A. A fusion approach using Dempster–Shafer theory and Bayesian learning

FDS of Dempster–Shafer theory and Bayesian learning

Dempster–Shafer theory and Bayesian learning is a hybrid approach for credit card fraud detection [2][5][12][15] which combines evidences from current as well as past behavior. Every cardholder has a certain type of shopping behavior,

978-1-4244-9394-4/11/$26.00 ©2011 IEEE

which establishes an activity profile for them. This approach proposes a fraud detection system using information fusion and Bayesian learning of so as to counter credit card fraud.

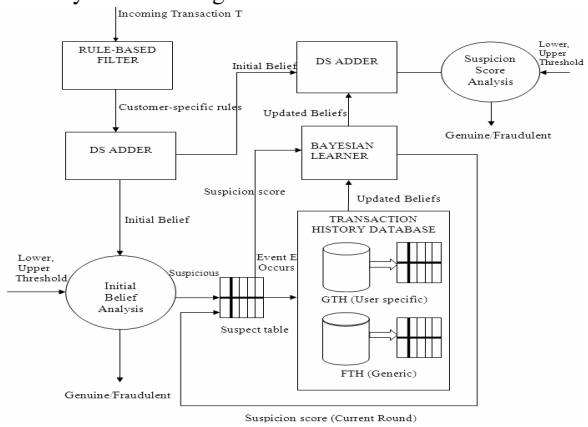

Figure 1. Block diagram of the proposed fraud detection system

The FDS system consists of four components, namely, rule-based filter, Dempster–Shafer adder, transaction history database and Bayesian learner. In the rule-based component, the suspicion level of each incoming transaction based on the extent of its deviation from good pattern is determined. Dempster–Shafer's theory is used to combine multiple such evidences and an initial belief is computed. Then the initial belief values are combined to obtain an overall belief by applying Dempster–Shafer theory. The transaction is classified as suspicious or suspicious depending on this initial belief. Once a transaction is found to be suspicious, belief is further strengthened or weakened according to its similarity with fraudulent or genuine transaction history using Bayesian learning.

It has high accuracy and high processing Speed. It improves detection rate and reduces false alarms and also it is applicable in E-Commerce. But it is highly expensive and its processing Speed is low.

B. BLAST-SSAHA Hybridization for Credit Card Fraud Detection

BLAST-SSAHA in credit card fraud detection

The Hybridization of BLAST and SSAHA algorithm [1][6][14] is refereed as BLAH-FDS algorithm. Sequence alignment becomes an efficient technique for analyzing the spending behavior of customers. BLAST and SSAHA are the efficient sequent alignment algorithms used for credit card fraud detection.

BLAH-FDS is a two-stage sequence alignment algorithm in which a profile analyzer (PA) determines the similarity of an incoming sequence of transactions on a given credit card with the genuine cardholder's past spending sequences. The unusual transactions traced by the profile analyzer are passed to a deviation analyzer (DA) for possible alignment with past fraudulent behavior. The final decision about the nature of a transaction is taken on the basis of the observations by these two analyzers.

BLAST-SSAHA Hybridization

When a transaction is carried out, the incoming sequence is merged into two sequences time-amount sequence TA. The TA is aligned with the sequences related to the credit card in CPD. This alignment process is done using BLAST.

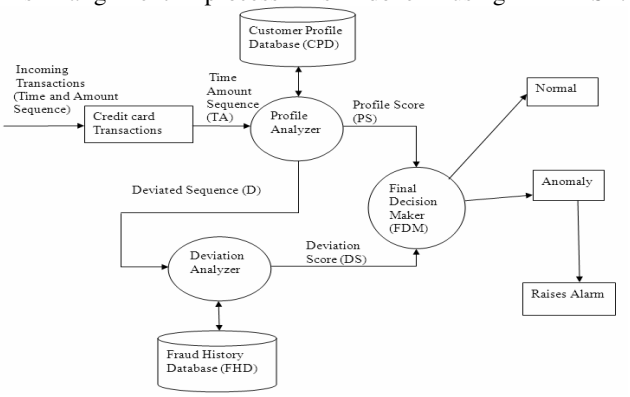

Figure 2. Architecture of BLAST and SSAHA Fraud Detection System

SSAHA algorithm [9] is used to improve the speed of the alignment process. If TA contains genuine transaction, then it would align well with the sequences in CPD. If there is any fraudulent transactions in TP, mismatches can occur in the alignment process. This mismatch produces a deviated sequence D which is aligned with FHD. A high similarity between deviated sequence D and FHD confirms the presence of fraudulent transactions. PA evaluates a Profile score (PS) according to the similarity between TA and CPD. DA evaluates a deviation score (DS) according to the similarity between D and FHD. The FDM finally raises an alarm if the total score (PS - DS) is below the alarm threshold (AT).

The performance of BLAHFDS is good and it results in high accuracy. At the same time, the processing speed is fast enough to enable on-line detection of credit card fraud. It Counter frauds in telecommunication and banking fraud detection. But it does not detect cloning of credit cards

C. Credit Card Fraud Detection using Hidden Markov Model

Hidden Markov Model

A Hidden Markov Model is a double embedded stochastic process with used to model much more complicated stochastic processes as compared to a traditional Markov model. If an incoming credit card transaction is not accepted by the trained Hidden Markov Model with sufficiently high probability, it is considered to be fraudulent transactions.

Use Of HMM For Credit Card Fraud Detection

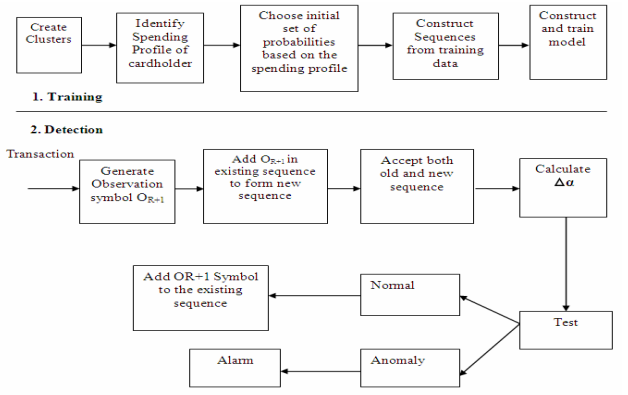

Figure 3. Process Flow of the Proposed FDS

A Hidden Markov Model [3] is initially trained with the normal behavior of a cardholder. Each incoming transaction is submitted to the FDS for verification. FDS receives the card details and the value of purchase to verify whether the transaction is genuine or not. If the FDS confirms the transaction to be malicious, it raises an alarm and the issuing bank declines the transaction. The concerned cardholder may then be contacted and alerted about the possibility that the card is compromised.

HMM never check the original user as it maintains a log. The log which is maintained will also be a proof for the bank for the transaction made. HMM reduces the tedious work of an employee in bank since it maintains a log. HMM produces high false alarm as well as high false positive.

D. Fuzzy Darwinian Detection of Credit Card Fraud

The Evolutionary-Fuzzy System

Fuzzy Darwinian Detection system [4] uses genetic programming to evolve fuzzy logic rules capable of classifying credit card transactions into “suspicious” and “non-suspicious” classes. It describes the use of an evolutionary-fuzzy system capable of classifying suspicious and non-suspicious credit card transactions.The system comprises of a Genetic Programming (GP) search algorithm and a fuzzy expert system.

Data is provided to the FDS system. The system first clusters the data into three groups namely low, medium and high. The GPThe genotypes and phenotypes of the GP System consist of rules which match the incoming sequence with the past sequence. Genetic Programming is used to evolve a series of variable-length fuzzy rules which characterize the differences between classes of data held in a database. The system is being developed with the specific aim of insurance-fraud detection which involves the challenging task of classifying data into the categories: "safe" and "suspicious". When the customer’s payment is not overdue or the number of overdue payment is less than three months, the transaction is considered as “non-suspicious”, otherwise it is considered as “suspicious”. The Fuzzy Darwinian detects suspicious and non -suspicious data and it easily detects stolen credit card Frauds.

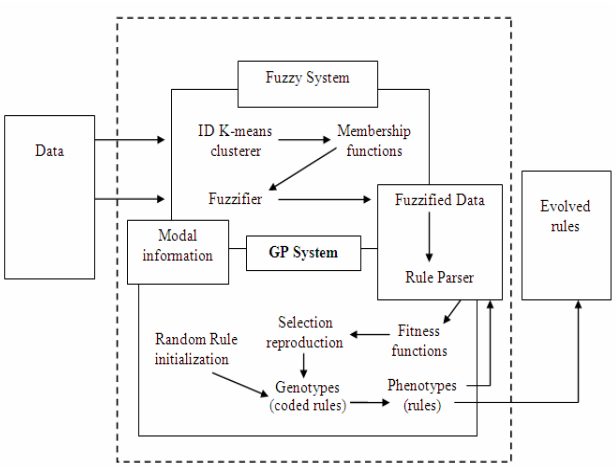

Figure 4. Block diagram of the Evolutionary-fuzzy system

The complete system is capable of attaining good accuracy and intelligibility levels for real data. It has very high accuracy and produces a low false alarm, but it is not applicable in online transactions and it is highly expensive. The processing speed of the system is low.

E. Credit Card Fraud Detection Using Bayesian and Neural Networks

The credit card fraud detection using Bayesian and Neural Networks are automatic credit card fraud detection system by means of machine learning approach. These two machine learning approaches are appropriate for reasoning under uncertainty.

An artificial neural network [5][7][9][10][16] consists of an interconnected group of artificial neurons and the commonly used neural networks for pattern classification is the feed-forward network. It consist of three layers namely input, hidden and output layers. The incoming sequence of transactions passes from input layer through hidden layer to the output layer. This is known as forward propagation. The ANN consists of training data which is compared with the incoming sequence of transactions. The neural network is initially trained with the normal behavior of a cardholder. The suspicious transactions are then propagated backwards through the neural network and classify the suspicious and non-suspicious transactions. Bayesian networks are also known as belief networks and it is a type of artificial intelligence programming that uses a variety of methods, including machine learning algorithms and data mining, to create layers of data, or belief. By using supervised learning, Bayesian networks are able to process data as needed, without experimentation. Bayesian belief networks are very effective for modeling situations where some information is already known and incoming data is uncertain or partially unavailable. This information or belief is used for pattern identification and data classification.

A neural network learns and does not need to be reprogrammed. Its processing speed is higher than BNN. Neural network needs high processing time for large neural

networks. Bayesian networks are supervised algorithms and they provide a good accuracy, but it needs training of data to operate and requires a high processing speed.

III. COMPARISON OF VARIOUS FRAUD DETECTION SYSTEMS PARAMETERS USED FOR COMPARISON

The Parameters used for comparison of various Fraud Detection Systems are Accuracy, Fraud Detection Rate in terms of True Positive and false positive, cost and training required, Supervised Learning. The comparison performed is shown in Table 1.

Accuracy: It represents the fraction of total number of transactions (both genuine and fraudulent) that have been detected correctly.

Method: It describes the methodology used to counter the credit card fraud. The efficient methods like sequence alignment, machine learning, neural networks are used to detect and counter frauds in credit card transactions.

True Positive (TP): It represents the fraction of fraudulent transactions correctly identified as fraudulent and genuine transactions correctly identified as genuine. False Positive (FP): It represents fraction of genuine transactions identified as fraudulent and fraudulent transactions identified as genuine.Training data: It consists of a set of training examples. The fraud detection systems are initially trained with the normal behavior of a cardholder.

Supervised Learning: It is the machine learning task of inferring a function from supervised training data. Comparison Results

The Comparison table was prepared in order to compare various credit card fraud detection mechanisms. All the techniques of credit card fraud detection described in the table 1 have its own strengths and weaknesses.

Results show that the fraud detection systems such as Fuzzy Darwinian, Dempster and Bayesian theory have very high accuracy in terms of TP and FP. At the same time, the processing speed is fast enough to enable on-line detection of credit card fraud in case of BLAH-FDS and ANN.

IV. CONCLUSION

Efficient credit card fraud detection system is an utmost requirement for any card issuing bank. Credit card fraud detection has drawn quite a lot of interest from the research community and a number of techniques have been proposed to counter credit fraud. The Fuzzy Darwinian fraud detection systems improve the system accuracy. Since The Fraud detection rate of Fuzzy Darwinian fraud detection systems in terms of true positive is 100% and shows good results in detecting fraudulent transactions. The neural network based CARDWATCH shows good accuracy in fraud detection and Processing Speed is also high, but it is limited to one-network per customer. The Fraud detection rate of Hidden Markov model is very low compare to other methods. The hybridized algorithm named BLAH-FDS identifies and detects fraudulent transactions using sequence alignment tool. The processing speed of BLAST-SSAHA is fast enough to enable on-line detection of credit card fraud. BLAH-FDS can be effectively used to counter frauds in other domains such as telecommunication and banking fraud detection. The ANN and BNN are used to detect cellular phone fraud, Network Intrusion. All the techniques of credit card fraud detection discussed in this survey paper have its own strengths and weaknesses. Such a survey will enable us to build a hybrid approach for identifying fraudulent credit card transactions.

Table 1 Comparison of various fraud detection method

Parameter	Fusion of Dempster Shafer theory and Bayesian learning	Hybridization of DLAST and SSAHA	HMM	Artificial Neural Networks and Bayesian Neural Networks		Fuzzy Darwinian detection
				ANN	BNN	
Study	Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar (2009)	Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar (2009)	Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K Majumdar (2008)	Chi Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, Bernard Manderick (1999)		Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-Uk Choi (2000)
Method	Machine Learning	Sequence Alignment	Hidden Markov Model	Artificial Intelligence, Machine Learning	Artificial Intelligence, Machine Learning	Genetic Programming, Fuzzy Logic
Fraud Detection TP%	98%	80%	70%	77%	34%	100%
Fraud Detection FP%	10%	10%	20%	13%	10%	5.79%
Processing Speed	Medium	Very High	High	High	Low	Low
Training required	Yes	No	Yes	Yes	Yes	Yes
Supervised Learning	Supervised	Unsupervised	Semi supervised	Supervised	Supervised	Supervised
Cost	Implementation is expensive	Inexpensive	Quite expensive	Expensive	Expensive	Highly Expensive
Accuracy	High	High	Medium	Medium	Medium	Very high
Research issues addressed	Intrusion detection in many database applications Applicable in E-Commerce	Applicable in telecommunication and banking fraud detection Online detection, cost is inexpensive	Applicable in online detection of credit card fraud. No need to check the original user as it maintains a log	Cellular phone fraud, Calling card fraud, Computer Network Intrusion Applicable in E-Commerce		Easily detect stolen credit card Frauds Detect suspicious, non-suspicious data
Research Challenges	Processing speed is very low	Cannot detect cloning of credit card fraud	High false alarm, False positive is high	Needs training to operate and requires high processing time for large neuralnetworks and BNN		Not applicable in E-commerce, Difficult to implement

REFERENCES

[1] Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection," *IEEE Transactions On Dependable And Secure Computing*, vol. 6, Issue no. 4, pp.309-315, October-December 2009.

[2] Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning," *Special Issue on Information Fusion in Computer Security*, Vol. 10, Issue no 4, pp.354-363, October 2009.

[3] Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar, "Credit Card Fraud Detection using Hidden Markov Model," *IEEE Transactions On Dependable And Secure Computing*, vol. 5, Issue no. 1, pp.37-48, January-March 2008.

[4] Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-Uk Choi, "Fuzzy Darwinian Detection of Credit Card Fraud," *In the 14th Annual Fall Symposium of the Korean Information Processing Society*, 14th October 2000.

[5] Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, Bernard Manderick, "Credit card fraud detection using Bayesian and neural networks," *Interactive image-guided neurosurgery*, pp.261-270, 1993.

[6] Amlan Kundu, S. Sural, A.K. Majumdar, "Two-Stage Credit Card Fraud Detection Using Sequence Alignment," *Lecture Notes in Computer Science, Springer Verlag, Proceedings of the International Conference on Information Systems Security*, Vol. 4332/2006, pp.260-275, 2006.

[7] Simon Haykin, "Neural Networks: A Comprehensive Foundation," 2nd Edition, pp.842, 1999.

[8] A. Chiu, C. Tsai, "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection," *Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service*, pp.177-181, 2004.

[9] R. Brause, T. Langsdorf, M. Hepp, "Neural Data Mining for Credit Card
Fraud Detection, *"International Conference on Tools with Artificial Intelligence*, pp.103-106, 1999.

[10] Ghosh, D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," *Proceedings of the International Conference on System Science*, pp.621-630, 1994.

[11] Z. Ning, A.J. Cox, J.C. Mullikin, "SSAHA: A Fast Search Method for Large DNA Databases," *Genome Research*, Vol. 11, No. 10, pp.1725-1729, 2001.

[12] Lam, Bacchus, "Learning bayesian belief networks: An approach basedon the MDL principle," *Computational Intelligence*, Vol. 10, Issue No. 3, pp.269–293, August 1994.

[13] Manoel Fernando Alonso Gadi, Xidi Wang, Alair Pereira do Lago, "Credit Card Fraud Detection with Artificial Immune System," *Lecture Notes in Computer Science*, Vol. 5132/2008, pp.119-131, 2008.

[14] Tom Madden, "The BLAST Sequence Analysis Tool", 2003.

[15] M. Mehdi, S. Zair, A. Anou and M. Bensebti," A Bayesian Networks in Intrusion Detection Systems," *International Journal of Computational Intelligence Research*, Issue No. 1, pp.0973-1873 Vol. 3, 2007.

[16] Ray-I Chang, Liang-Bin Lai, Wen-De Su, Jen-Chieh Wang, Jen-Shiang Kouh, "Intrusion Detection by Backpropagation Neural Networks with Sample-Query and Attribute-Query," *Research IndiaPublications,* pp.6-10, November 26, 2006.

[17] C. Phua, V. Lee, K. Smith, R. Gayler, "A Comprehensive Survey of Data Mining-based Fraud Detection Research," *Artificial Intelligence Review*, 2005.

[18] E. Aleskerov, B. Freisleben, B. Rao, "CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection," *Proceedings of IEEE/IAFE Conference on Computational Intelligence for Financial Engineering (CIFEr)*, pp.220-226, 1997.

[19] Philip K. Chan ,Wei Fan, Andreas L. Prodromidis, Salvatore J. Stolfo, "Distributed Data Mining in Credit Card Fraud Detection," *IEEE Intelligent Systems ISSN,* Vol. 14 , Issue No. 6, Pages: 67 – 74, November 1999.

[20] Barry Masuda, "Credit Card Fraud Prevention: A Successful Retail Strategy," *crime prevention*, Vol. 6, 1986.

[21] Linda Delamaire, Hussein Abdou, John Pointon, "Credit card fraud and detection techniques: a review," *Banks and Bank Systems,* pp. 57-68*,* 2009.

[22] Tej Paul Bhatla, Vikram Prabhu & Amit Dua "Understanding Credit Card Frauds," 2003.

[23] Russell, Norvig ," Artificial Intelligence – A Modern Approach," 2nd Edition, 2003.

[24] S.F.Altschul, W. Gish, W. Miller, W. Myers, J. Lipman, "Basic Local Alignment Search Tool," *Journal of Molecular Biology*, Vol. 215, pp.403-410, 1990.

[25] Ezawa.K. & Norton.S,"Constructing Bayesian Networks to Predict Uncollectible Telecommunications Accounts," IEEE *Expert*, October; 45-51, 1996.

[26] Fan, W. Miller, M.Stolfo, S.Lee & P Chan, "Using Artificial Anomalies to Detect Unknown and Known Network Intrusions," *Proc. of ICDM01,* pp.504-507, 2001.