World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 1, No.8, 326- 332, 2011

Detect CNP Fraudulent Transactions

Adnan M. Al-Khatib
CIS Dept., College of IT and Computer Science,
Jerash University, Jerash, Jordan

Abstract— Money in the e-commerce network, represents information moving at the speed of light, where fraud (digital crime) within the banking and financial services happened very fast and can cost billions of dollars each year-undetected and unreported. In this paper I present a comprehensive framework that mines and detect fraudulent transactions of Card-Not-Present (CNP) in the e-payment systems with a high degree of accuracy.

Keywords- Credit Card; Card-Not-Present; Fraud Detection; Data Mining; Profiling; Accuracy; Rule.

I. INTRODUCTION

Fraud can be described as "a dishonest actions to make false statements in order to gain money or benefit from an individual or from an organization". Fraud on the Internet includes several types such as theft of funds through illegal transfers, theft of credit card details, illegal credit card use, and others [7].

Fraud detection is a process that can use data mining techniques to detect fraudulent transaction. In this paper, I propose a comprehensive framework that mines fraudulent transactions of Card-Not-Present (CNP) in the e-payment systems with high degree of accuracy. Our research used the user account profiling techniques to discover fraudulent transactions in CNP payment systems. To detect fraud using profiling technique, it is necessary to determine the normal behavior of each user account with respect to certain indicators, and to determine when that behavior has deviated significantly.

This paper presents an overview of the proposed system in section 1. Section 2 gives an overview of related works. Section 3 gives an overview of fraudulent activities in financial area. Section 4 describes our proposed system. Section 5 presents and evaluates the results. Section 6 discusses the method and compares it with other techniques. Section 7 concludes the research.

II. RELATED WORKS

In my survey [1], I presented an evaluation and comparisons for some detection techniques such as neural network (NN0, Rule Induction (RI), Expert systems (ES), Case-based reasoning (CBR), Genetic Algorithms (GA), Inductive logic programming (ILP) and Regression. Our study shows that the efficiency and performance of these techniques depend on there capability in dealing with several problems such as: noisy and missing in the data used, the performance measure used, scalability, different data types used, explanation capability of the technique, ease of integration with other systems, ease of operation, and skewed distribution of the data used [3, 6, 12, 13]. A comparison of these techniques with our method shows that our method outperformed all of them.

III. FINANCIAL CRIMES

Financial Crimes consists several types of fraud such as: Credit-Card Fraud, Card-Not-Present (Internet credit-card) Fraud, Loan Default, and Bank Fraud [8]. Credit-Card Fraud can be a result of a stolen card with the PIN number, or as a result of the theft of an individual's identification (Social security number and home address) in order to create a new account under false or stolen identities. Credit-card theft will defraud the card issuer or merchant. Card-Not-Present Fraud like Internet and phone-order sales transactions. They are also time-sensitive crimes. In this type of fraud, thieves leave characteristic footprints. For example, fraud rates increase at certain time of the day, and order coming from certain countries exhibit a higher percentage of fraud. Loan Default fraud involves the manipulation and inflation of an individual credit rating prior to performing a "sting", leading to a loan default and a loss for the financial service provider. Bank Fraud involves the creation of fictitious bank account for the conduit of money and the siphoning of other legitimate accounts.

The critical factors for detecting all of these financial fraud crimes is to know the behavior of credit, bank accounts, and loan accounts and developing an understanding of the categories of customers. Data mining can be used to spot outliers or account usage that are normal and out of character.

IV. RESEARCH MOTIVATION

Our research purpose is "to present a high accuracy method or prototype to detect Card-Not-Present (CNP) Fraudulent transactions in the e-payment systems by integrating data from multiple databases (e.g., bank transactions, federal/state crime history DBs); and then using suitable and effective data mining and artificial intelligence (AI) tools to find unusual access sequences". Accuracy means high detection rate (percentage of fraudulent transactions that are detected) and low false positive rate (percentage of normal transactions that is falsely determines to be fraudulent) [4, 12].

Researchers have developed two general categories of detection techniques; misuse and anomaly detections. In misuse detection, well-known fraudulent transactions are encoded into patterns, which are then used to match new transactions to identify the fraudulent ones. In anomaly detection, normal behavior of user are first summarized into normal profiles, and then used as yardsticks, so that run-time activities that result in significant deviation from the user profiles are considered as probable fraudulent transactions. In my research I am going to use the user account profiling techniques to discover fraudulent transactions in CNP payment systems. To use this technique three issues arise [14]:

1 - Which transaction features are important? Which features or combinations of features are useful for distinguishing legitimate behavior from fraudulent behavior?
2 - How should profiles be created? Given an important feature identified in step 1, how should we characterize the behavior of a subscriber with respect to the feature?
3 - When should alarms be issued? Given a set of profiling criteria identified in step 2, how should we combine them to determine when fraud has occurred?

A. Proposed System

Based on [14] I created my Proposed System (figure 1). The system consists of a warehouse, profiling module, software programs, a profiler monitor construction module, a storage component and a data mining classifiers modules such as artificial neural network (ANN), Find laws and others.

The warehouse contains transactional and historical data. Technical analysis (TA) tools (e.g., Link analysis, decision tree, etc) will analyze this data to generate the fraudulent rules as IF/THEN rules. The software programs will be used to normalize, manipulate and store the rules that are generated and selected by the analysis tools and a machine learning selection program. The profiler monitor construction module generates a set of profilers from the discovered fraud rules and a set of profiling templates instantiated by the fraud rule conditions. In the storage component, data for different purposes will be stored such as the user profiles, training and testing data for training the detector module and so on. The training data is provided as

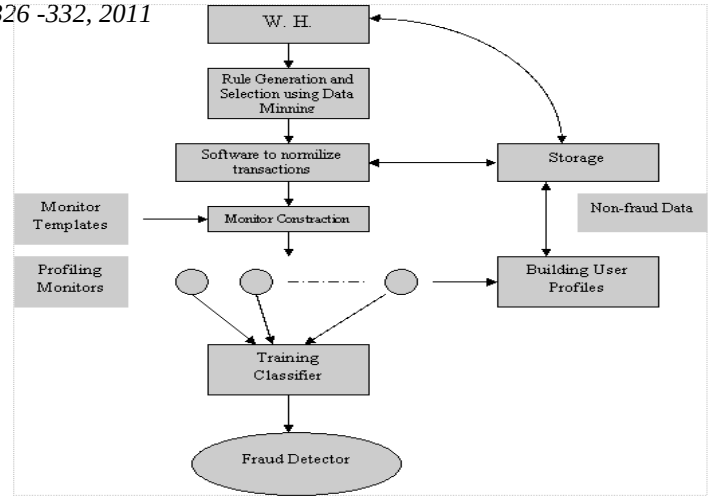

Figure 1: Proposed System.

input to the detector module, where a classifier is trained to classify the different observations of patterns into good or bad occurrences. Profiling module will be used to build user profiles according to the fraud rules conditions that selected and the set of templates provided. Each user will have a profile represent the normal profiling behaviors for that user account such as: cardholder shopping habits, frequency of purchases, average purchases, location of purchases, and other transactional factors. In the detector use step these profiles are used to match new transaction with the user profile, decide if there is a significant deviation of the transaction from the user profile, and give a numeric values representing the fraudulent activity for each fraudulent indicator. All these numeric values then passed to the detector as input parameters, where the detector combines these evidence factors to produce an output recommendation about the new transaction.

B. Detector Construction

Detector construction process consists three stages: The first stage is the data mining analysis, involves combing through the transactional data searching for indicators of fraud with certainty factors above a user threshold. The transactional data are organized by account, and each transactional record is labeled as fraudulent or legitimate. When the rule learning (RL) program is applied to an account's transactions it produces a set of rules that serve to distinguish, within that account, the fraudulent transactions from the legitimate transactions. As an example, the following rule would be a relatively good indicator of fraud:

(TIME-OF-DAY = NIGHT) AND (LOCATION = far location from customer zip-code) ==➔ FRAUD, with Certainty factor = 0.85

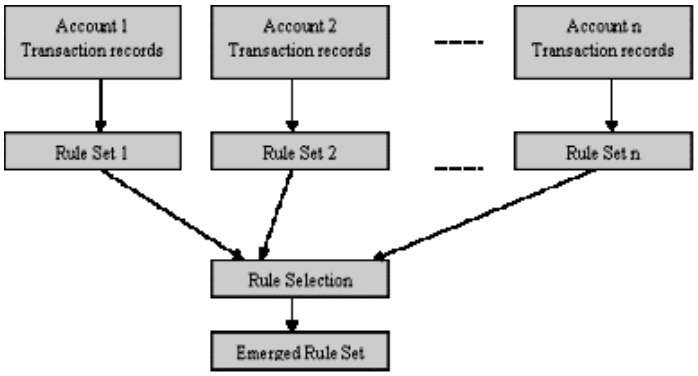

Figure 2: First stage, rule generation and selection.

This rule denotes that a transaction placed at night from a far location from the customer home address is likely to be fraudulent. The certainty factor = 0.85 means that, for this account, a transaction matching this rule has an 85% probability of being fraudulent.

Each account generates a set of rules. After all accounts have been processed, a rule selection step is performed to derive a general covering set of rules that will serve as fraud indicators, figure 2.

In the second stage, the profiler constructor is given a set of rules and a set of templates, figure 4, and generates a profiler from each rule-template pair. Every profiler has a training step, figure 3, in which it is trained on typical (non-fraud) account activity; and a Use step, figure 4, in which it describes how far from the typical behavior a current account transaction is.

The third stage of detector construction learns how to combine evidence from the set of profilers generated by the previous stage. In training, the profilers' outputs are presented along with the desired output to the classifier. The evidence combination learns which combinations of profiler outputs indicate fraud

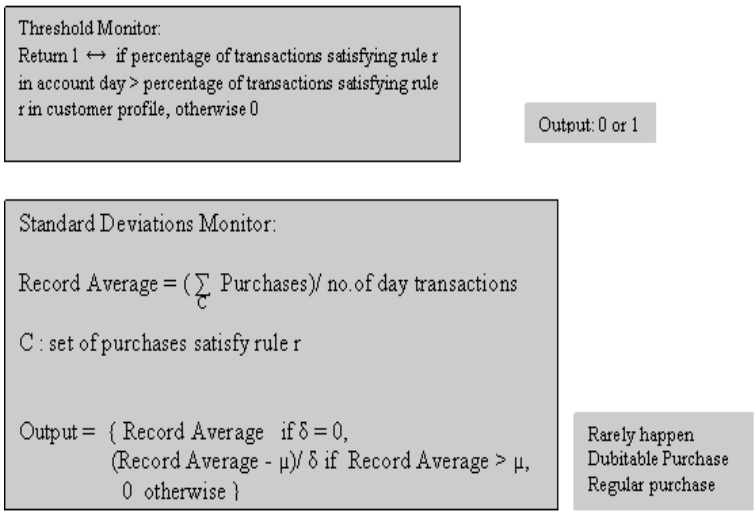

Figure 3: Profiling step

with high confidence. A feature selection process is used to reduce the number of profilers in the final detector. This simplifies the final detector and increases its accuracy and performance. The final output of the constructor is a detector that profiles each user's behavior based on several indicators, and produces an alarm if there is sufficient evidence of fraudulent activity.

New transactions are matched first with customer profiles, checking for deviation from normal behavior, and assign numeric values for the profilers monitors. These profilers monitors output then used as input for the detector,

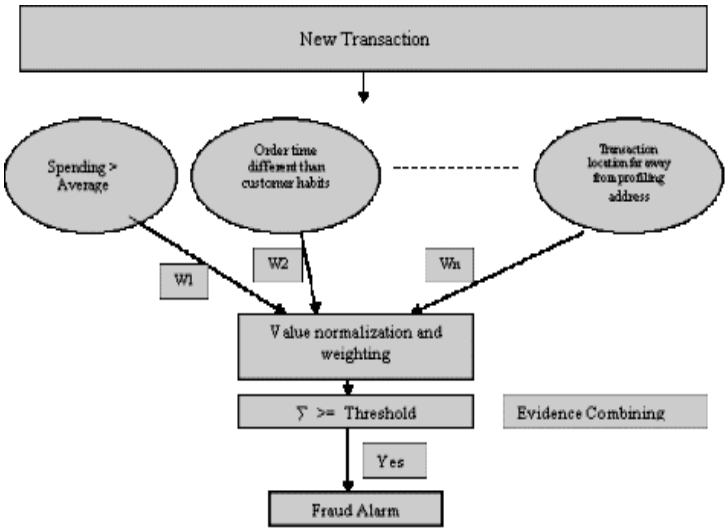

Figure 4: Use step.

which are combined and the detector produce an alarm if there is sufficient evidence of fraudulent activity. Figure 5 is an example of evaluating a new transaction.

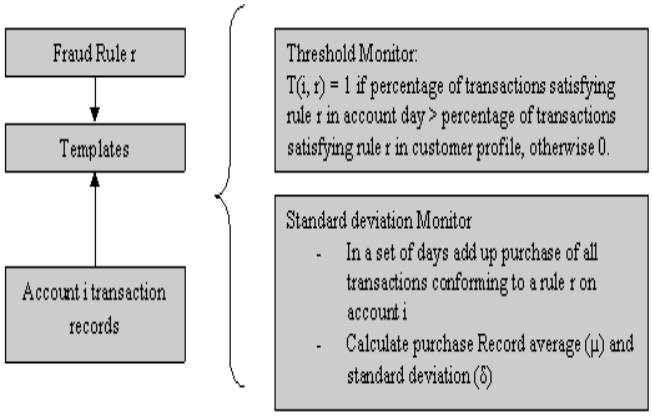

Figure 5: Detector usage.

V. EXPERIMENT AND EVALUATION

A. Data

Data for our study can be collected from different available resources, but because of security considerations of this type of data from its financial resources, I used hypothetical data, which is generated by a simulation programs that generate the data randomly.

Based on reference [5], I wrote tow simulation Visual Basic programs. First program used to create randomly customer database. The second program used to generate randomly customer transactions dataset. Customer database created with 105 customer records each record contains 26 fields contains basic information about the customer such as customer account number, customer first name, customer last name, customer address, customer email, etc. The second program generates randomly 11282 transactions, 4850 transactions of them are fraudulent transactions and the rest (6432) are legitimate transactions. Each transaction contains 53 attributes. In addition, a normalization and profiling step performed on the dataset to derive new calculated attributes with 0 or 1 value. For example "dif_cus_email" attribute hold the value 1 if customer email address in the transaction is different from the customer email address in the customer database and hold the value 0 otherwise.

B. Data Mining Software

PolyAnalyst is a powerful multi-strategy data mining system that implements a broad variety of methods for the automatic data analysis [11]. It contains eleven advanced knowledge discovery algorithms such as Classification, Decision Tree, Find Dependencies, Find Laws, and NN Predictor. It can perform a thorough analysis of data, automatically extracting the precious knowledge from an investigated database and presenting it in symbolic form easily understood by a human.

C. Learning and selection Fraud Rules

An analysis step performed on each account using the Decision tree algorithm to generate indicators of fraud in the form of classification rules. 512 rules were generated and summarized in an Excel worksheet. The following are some examples of these rules:

112 "If dif_cus_email And trans_time_of_day = Night" ➔ Fraud with CF = 86.7%"
161 "If dif_bill_ship_last_name And trans_amount >=100 And trans_time_of_day = Night" ➔ Fraud with CF = 87%"

The first rule for account 112 says that if customer email address in the transaction is different from email address in the customer profile and transaction time is night then the transaction is fraudulent with certainty factor 86.7%.

These 512 rules are normalized and sorted and a selection process performed on them to select the most general rules that covered two or more accounts, and only 46 rules were selected to build the profiling monitors.

D. Constructing Profiling Monitors

The monitors were used to build customer profiles; each monitor has a Profiling step and a Use step. In the profiling step, the monitor is applied to a segment of an account's typical (non-fraud) transactions in order to measure the account's normal activity. Statistics from this profiling period are saved with the account. The following are these statistics information calculated using Excel functions:

Percentage of attribute True values from the transactions (profile Threshold))
Standard Deviation for amounts for attribute True values in transactions (σ) (profile)
Record Average of amount for attribute True values in Transactions satisfy rule conditions (profile) (μ)

In the monitor's Use phase, the monitor processes a single account-day at a time. The monitor references the normally measures calculated in profiling, and generates a numeric value describing how abnormal the current account day is. The following are the statistical information calculated in the Use step by using Excel functions:

Number of attribute True values in account day

Number of Transactions in account day

Percentage of attribute True values in account day

Sum of amount for attribute True values in Transactions

Record Average of amount for attribute True values in Transactions satisfy rule

Then these information compared with the information calculated in the profiling step using two templates (Threshold and standard deviations) mentioned in the methodology section to calculate two output monitors for each rule. So 92 output monitors created for each account days. The following is an example of these 92 output monitors created for account day of one account:

101	1	YES		(trans_acct_no,	fraud_flag,
fraud_yes_no)					
0	1	0	0	0	0	0
	1	0	0	0	0	1
	1	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	1		(46 Threshold monitors)	

0 424.3333333 0 0 0 0
0 2.210193235 0 0.128199211
0 0 0.229560816 0.379616247
0 0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0
0 0 0 0 0 0
0 115.9166667 (46 standard deviation monitors)

These output monitors were calculated for several account days for each account over a period of four months, and calculated for 48 accounts. This process produced 2083 output monitors records for the 48 accounts. These 2083 records divided into training set and testing set to be used for detector training and testing processes.

E. Combining Evidence from the Monitors

The training monitors set used as input to three classifier programs in order to perform Evidence combination and build classification rules and Detector. A selection process was performed on the tanning and testing monitors sets in order to reduce the number of independent attributes. This process reduced the monitors to 19 monitors only. The other monitors or rules do not perform well when used in monitors, and some monitors overlap in their fraud detection coverage. In addition, this process chooses a small set of useful monitors which simplifies the final detector and increases its accuracy and performance.

F. Evaluation

The three classification algorithms produce classification efficiency 100%. Summary statistics for testing the Detector rules is shown in table 1. The table shows that 192 monitors' records (90 fraud and 102 non-fraud) used for testing, and the three classifiers produced the same results with overall accuracy reached (100% - ((107-90)*100/192)) = 91.2%.

Accuracy means that the model will have high detection rate (percentage of fraudulent
transactions that are detected) and low false positive rate (percentage of normal transactions that the system is falsely determines to be fraudulent) [3].

From the information in Table 1 we can calculate the accuracy as follows:

Over all accuracy = 100% - percentage of incorrectly classified records

= 100%- (17/192)* 100 = 91.2%.

TABLE 1: TESTING SUMMARY STATISTICS

(yes/no) attributes:	Values	N of 1	N of 0
fraud_yes_no	192	90 (46%)	102 (53%)
CL_PN_Liberal_Data	192	107 (55%)	85 (44%)
CL_LR_Liberal_Data	192	107 (55%)	85 (44%)
CL_FL_Liberal_Data	192	107 (55%)	85 (44%)

Detection Rate = percentage of fraudulent transactions that are detected = 100% (no fraudulent monitor records classified in the wrong class, which means No False negative error).

False alarm error = (17/192)*100% = 8.8% (legitimate records classified as fraudulent records)

The cost of false alarm can be estimated by the cost of a fraud analyst's time, which is very low and can be neglected comparing with the system benefits.

The training process repeated with different distribution of the training monitors set and the same testing monitors set and produced the same testing results with errors in the same accounts. This shows us that part of the error percentage (8.8%) refers to human errors in some accounts data calculations and to some transactions with very low spending amounts neglected from verification because verification cost is greater than transaction coast (according to financial institute policy). I removed these accounts from the monitor's sets and repeated the training and testing process and I get an overall accuracy results reached to 99.2%, detection rate equal 100%, and false alarm equal 0.8%. The summary statistical results for this case are shown in table 2:

TABLE 2: TESTING SUMMARY STATISTICS

(yes/no) attributes:	Values	N of 1	N of 0
fraud_yes_no	120	53 (44%)	67 (55%)
CL_PN_L_Data_Rules	120	54 (45%)	66 (55%)
CL_FL_L_Data_Rules	120	54 (45%)	66 (55%)
MBE_L_Data_Rules	120	54 (45%)	66 (55%)
CL_LR_L_Data_Rules	120	54 (45%)	66 (55%)

From the above statistical results and testing processes we notice that all the error percentage represent False positive errors (false alarm) corresponds to wrongly deciding that an account has been frauded. While there is no False negative error corresponds to letting frauded account-day go undetected. The cost of false alarm can be estimated by the cost of a fraud analyst's time, which is very low and can be neglected comparing with the system benefits.

VI. DISCUSSION

It is difficult to evaluate our method of fraud detection against existing fraud detection systems. Fraud detection departments and vendors of fraud detection systems protect details of their systems operations for trade and security purposes. I evaluated my method against known fraud

techniques and against a collection of techniques as I presented them in our survey [1].

In our detection method, account context is important in the rule learning step: a global transaction set taken from all accounts and applying a rule learning algorithm to this set would lose information about each account's normal behavior. For example, a transaction from the east area for a customer who lives in the west area at night would be a fraud transaction, while it would be legitimate transaction for a customer who lives in or near the east area. Standard detection technique usually not consider account context in the detection process.

Applying standard classification algorithms to the account data is difficult for several reasons. For example, the description language is very detailed because many thousands of attribute values appear in the data and standard simple classifier will not perform well. In my method I performed a profiling step to change all attributes to logical (0/1) attributes which allow the classifier to perform very well.

Some fraud analysts believe that credit card fraud accompanied by large jumps in account usage or spending amount exceed normal average, and sophisticated mining of fraud is probably unnecessary. This is not true, for example in special occasion's time like Christmas, New Year and other occasion's usage of credit cards and spending amounts increased by customers. So a need for sophisticated mining system depending on combination of several variables or evidences is necessary.

Our method framework has three main components, and is more complex than other approaches, but each component has its own important contribution:

- Learning Component uncover specific indicators of fraudulent transactions.
- Our method shows the value of rule generation step, which does preserve account context.
- Our method shows the benefit of combining evidence from multiple monitors.
- Evidence combination step allows catching different and most of fraudulent transactions instead of depending only on a single or few monitors as happen in standard detection techniques.
- Composite of the three components in our system outperformed standard detector in which a significant piece of data or information is missing.

In addition my method can use different classifiers from simple linear to complex ones to combine evidence from the monitors and perform high accuracy.

VII. Conclusion

In this paper, I present a comprehensive framework that mines fraudulent transactions of Card-Not-Present (CNP) in the e-payment systems with high percentage of accuracy. The framework uses data mining analysis tools to discover indicators of fraudulent behavior by analyzing a massive amount of data, and then builds modules to profile each user account's normal behavior with respect to these indicators. The profilers capture the typical behavior of a user account and, in use, describe how far an account transaction is from this typical behavior. The profilers are combined into a single detector, which learns how to detect fraud effectively based on the profiler outputs. When the detector has enough evidence of fraudulent activity on an account transaction, based on the indications of the profilers, it generates an alarm.

Our detection method present high accuracy in predicting fraudulent transactions that exceeds 91.2% and in some cases reached 99.2% as overall accuracy, with fraud detection rate equal 100%. We noticed also that part of the error percentage refer to human error in entering and calculating the data. Our method can be implemented easily to automate all its steps and to produce an automated fraud detection system that can produce a very high percentage of accuracy.

Our detection framework is not specific to credit card fraud; it can be applied to several other fraud problems in different domains.

References:

[1] Adnan M. Al-Khatib and Ezz Hattab; "Credit Card Fraud Detection Techniques: A survey"; the 7th international conference (iiWAS2005); 2005; Vol 1; P.P. 505 – 516.

[2] Adnan M. Al-Khatib and Ezz Hattab; "Mining Fraudulent Transactions in e-payment Systems"; the 9th international conference (iiWAS2007); 2007; P.P. 179 – 189.

[3] Adnan M. Al-Khatib; "Mining Fraudulent Behavior in e-payment Systems"; Ph.D. Dissertation; 2007.

[4] Andreas L. Prodromidis; "Agent-Based Distributed Learning Applied to Fraud Detection"; Columbia University; 2000.

[5] "Advanced Integration Method (AIM) Implementation Guide Card-Not-Present Transactions", Version 1.0, 2005, Merchant Commerce and Payment Services.

[6] Clifton Phua; "Minority Report in Fraud Detection: Classification of Skewed Data"; Sigkdd Explorations, Vol. 6.

[7] Commonwealth of Australia; "The changing nature of fraud in Australia"; 2000.

[8] Jesus Mena; "Investigative Data mining for Security and Criminal Detection"; B. H. pub. Company; 2003.

[9] Jussi Ahola and Esa Rinta-Runsala; "Data mining case studies in customer profiling"; Research report TTE1-2001-29; VTT Information Technology; 2001
[10] Margaret H. Dunham; "Data Mining Introductory and Advanced Topics"; Prentice Hall; 2003
[11] "PolyAnalyst 4 user Manual"; 2002 Megaputer Intelligence, Inc.
[12] Salvatore J. Stolfo; "Credit Card Fraud Detection Using Meta-Learning "; Columbia University; 1997.
[13] Salvatore J. Stolfo and Wei Fan "Cost-based Modeling for Fraud and Intrusion Detection: Results from the JAM Project"; Columbia University; 0-7695-0490-6/99, 1999 IEEE.
[14] Tom Fawcett and Foster Provost; "Adaptive Fraud Detection"; Data Mining and Knowledge Discovery; 1997.