A Novel Approach for Behavior based Charge Card Fraud Detection using Support Vector Machines

Ms. D. Manjula[1] Ms. J. Thilagavathi[2]
[1]Research Scholar [2]Assistant Professor
[1,2]Department of Computer Science
[1,2]NGM College of Arts and Science, Pollachi

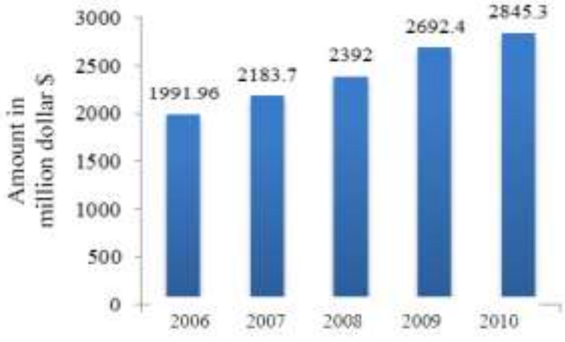
Abstract— Online payments are provided with extreme safety, even though, it is not unrestricted from faults. Decline in virtual extortions is the essential for the current day. One of the primary e-transactions that face these threats are the credit/debit cards. This method habits bundling and outlier discovery for discovery the illicit communications. Primarily, the communications are clustered translation to the attribute measured as obligatory for the finding process. Then each gathering is useful with the outlier discovery method for the discovery of data that departs from the existing data flow. This data has the highest possibility of being the fake data. The next procedure uses SVM, a binary classifier for healthier location. Then these methods are disposed to grips high level of fabricated positives. To decline these lying positives, multi-clustering is used. Hence, this suggest the manipulator with an in effect, mutual outlier finding appliance that reduces faults and offers more precise results for the user.

Key words: E-Payment, E-Commerce, Fraud deduction, cluster analysis

I. INTRODUCTION

The process of electronic payment has been used since 1970's. Many methods were designed to provide payments. After the internet came into being and after the incorporation of internet into many areas, the world has seen an explosive number of people using it. Payment using electronic means has started becoming a common mode of transaction during the late 90's. It was during this period when ideas came about for cashless transactions. These served as the basis for the electronic commerce or cashless transactions. This was also the period when electronic commerce came into full view. This got researchers working on the process and many academic studies conducted under this area. Since this process also involved a lot of commercial interest, many commercial agencies were also interested in this area and hence this area of research started bustling with activity. Many ideas were proposed for carrying out these cashless transactions, some of them were even launched into the market. But many failed to reach the targeted audience and hence failed. One of such methods, The Cyber cash launched payment systems. These systems achieved quite extensive deployment but failed to generate an economic return. At the same time many companies started up new methods of payments for B2C sector [1].

The Electronic Payment (e-payment) is a method of value exchange in electronic commerce, where the value is transferred via the Internet and communication technologies. The electronic payment systems have evolved from traditional payment systems and consequently the two types of systems have much in common. An electronic payment system denotes any kind of network service that includes the exchange of money for goods or services. E-payment is conducted in different e-commerce categories such as Business-to Business (B2B), Business-to-Consumer (B2C), Consumer-to- Business (C2B) and Consumer-to-Consumer (C2C) [2].The advent of e-commerce has initiated a new issue. It deals with the security while performing these transactions. One of the main problems that persuades in the issue of online transactions is the credit card frauds. These frauds have always been the highest of priority to the banks. These issues are of highest importance and are set to grow more. This problem is also set to strike the customer base as one of the highest priority issues. Over the recent years, these type of frauds has seen a huge raise, due to the increase in the usage of online transactions. The usage of payment cards has triggered a raise in the corresponding frauds and hence has become an issue of importance. The increase in usage of cards and crime related to it has converted these types of frauds into an organized crime activity. The use of credit cards has become one of the common tasks in the life of an average person. The usage of credit cards is one of the easier modes of transactions; hence it is also the most flexible and the easy way of payment. Detection of frauds performed using credit cards is a tedious as well as a crucial process. Since it also involves the customer's interests and flexibility, this process should be performed carefully. This has recently become a study of great importance due to the increase in the usage of credit cards and increase in the frauds detected. Even though this process is of high priority, handling this in an efficient way is very important, since a large customer base is involved. A simple flaw in the detection process might prove to be a heavy loss.



.

Fig. 1: Yearly Credit Card Fraud

As discussed in [3], in USA, the online retail sales were reported to be $ 144 billion in 2004, which was a 26% increase over 2003 [4]. It is also estimated that 87% of purchases made over the Internet are paid by credit card [5]. The Association of Payment and Clearing Services (APACS) report showed that the charge of charge card scam reached $ 966.74 million in 2004, which was an increase of 20% as compared to 2003 [6]. Additional review of over

160 concerns exposed that online fraud (committed over the Web or phone shopping) is 12 times difficult than offline fraud (committed by using a stolen physical card) [7].

Legacy systems, that consider account attributes like address of delivery, place of purchase, product purchased, purchase amount etc might prove to be misleading. Since the type of fraudulent transactions varies from time to time, and buying behavior of the users vary over time. Static rule based systems will be able to discover only the predefined type of frauds and not the new type of frauds, while the crime based activity of credit card frauds keeps growing. Confirming every suspected transaction from the genuine charge cardholder is not constantly possible or even practical due to the cost factor involved. There are many ways in which a credit card fraud takes place. In general these frauds can be categorized into two broad classifications:

A. Physical Card:

This type of fraud takes place when the actual card is in possession of the adversary. The cardholder either misplaces the card or his card is stolen and is then used by somebody else. This is the most common type of fraud that occurs. If the actual cardholder does not realize that his card is missing, then he would experience a large amount of loss. When a fraudster gets hold of such a card, he would attempt to perform large amount of transactions and purchases with very high value. These also take place in a very short span of time. Hence it shows a large deflection in the purchase pattern and the purchase volume when compared to the actual transactions carried out by the actual cardholder. Hence this can be easily detected with a simple analysis system.

B. Virtual Card:

This type of fraud is more difficult to detect, since the actual physical card is in possession of the cardholder. This type of fraud occurs when the fraudster gets hold of the card’s credentials instead of the actual physical card. This is possible when the user provides the credentials online in an unsecure payment site or in a spoofed site.

This can also occur when the fraudster possess counterfeit cards. This process is much more difficult to detect, since the actual user is not aware of another person using his card. If the fraudster makes purchases that are more similar to the actual user, then even an analysis system would not be able to detect the malicious transaction. This type of fraud comes to light only after the examination of the monthly statement by the actual cardholder.

One of the major problems facing the fraud detection is the lack of the availability of information regarding the credit card details and the frauds associated with them[8]. The details about credit card transactions are a set of vulnerable data, hence not many banks are willing to offer these information publicly. Even when such kind of information is offered, some significant parameters are eliminated from the dataset. Further, some parameters are substituted with dummy values for maintaining the confidentiality of the data. Only a very few publications provide valuable contributions in this field. Even these publications do not perform their experiments using the absolute data, instead, these experiments were performed on artificial data generated, that closely resembles the original data or real time datasets with dummy data as significant parameters. Hence this process does not prove to be much useful.

II. Review of Literature

Fraud detection is a critical part of the measures implemented for maintaining an attack tolerant database system. Though database management systems can provide intrusion prevention up to a certain extent by virtue of traditional access control mechanisms, they would not be sufficient for protection against syntactically correct but semantically damaging transactions [1]. Chung et al bring out that misuse detection in database systems has not been adequately addressed and propose DEMIDS, which can derive user profiles from database audit logs [2]. Lee et al suggest tagging the data objects with ”time semantics” and monitor behavior at the level of sensor transactions [3]. Hu and Panda concentrates on analyzing the dependencies among data items in a database [4].The field of Game theory has been explored for problems ranging from auctions to chess and its application to the domain of Information Warfare seems promising. Samuel et al bring out the role of Game theory in Information Warfare [5]. They highlight that one can utilize well-developed Game theory algorithms to predict future attacks and the differences and challenges in this domain as compared to traditional games like chess, such as limited examples, multiple simultaneous moves and no time constraints [6]. Liu and Li have presented a game-theoretic attack prediction model for attacks on IDS-protected systems [7].Credit card fraud detection has drawn lot of interest and a number of techniques, with special emphasis on data mining and neural networks, have been proposed to counter fraud in this field. Low et al described a method to implement a credit card system that would protect person’s identity using simple cryptographic blocks [14]. Ghosh and Reilly carried out a feasibility study for Mellon 266

V. Vatsa, S. Sural, and A.K. Majumdar Bank to determine the effectiveness of neural network for credit card fraud detection [15]. The neural network used for this study is the P-RCE (Restricted Coulomb Energy) neural network. The authors concluded that it was possible to achieve a reduction of 20% to 40% in the total fraud losses. Aleskerov et al presented CARDWATCH, a database mining system based on a neural network learning module [16] . The system trains a neural network with the past data of a particular customer, which can then be used to process the current spending behavior and detect anomalies and they assume that since the normal behavior of the thief is to purchase as much as possible in limited time, the anomaly in transactions will most probably be detected.

III. Methodology

Even though using SVM presents promising ways to fraud detection, the number of attributes provided to it is huge. This generally leads to large processing times. Some attributes might even have the probability of creating a negative impact on the result. Hence identification of attributes plays a crucial role in determining the accuracy of the final result. Further processing attributes that are dormant in the decision making process leads to unnecessary wastage of the processing time. All these problems can be overcome by selecting the appropriate attributes initially and then passing them to the SVM.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant formerly the input data will be transmuted into a condensed demonstration set of features. Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. Feature extraction is special form dimensionality reduction. Here, the features represent the relevant characteristics of the input data are chosen. Instead of using full size input one may use this reduced representation set. If it is properly chosen then it will give successful task. Best results are achieved when an expert constructs a set of application-dependent features. Nevertheless, if no such expert knowledge is available general dimensionality reduction techniques may help.

In this current process, feature extraction is performed as the initial process. The Principal Component Analysis (PCA) is used for feature compression. Principal component analysis (PCA) [wikipedia] is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components. Principal components are guaranteed to be independent if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables.

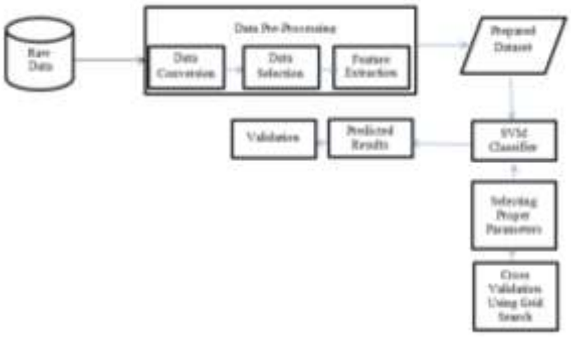


Fig. 2: Data processing in behavior based SVM

Principal component analysis (PCA) is a suitable tool for feature compression. The original feature space is reduced to low dimensional spaces but it will not affect the solution. The computational cost also less for training and testing the SVM because of using PCA.

PCA can be done by eigen value decomposition of a data covariance matrix, usually after mean centering the data for each feature. The covariance matrix [24] is calculated by the following,

$$Covariance(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)(y_i - y)$$

Equation No----------- (1)

where, x_i and y_i are the values of variables, x and y are the mean variables, n is number of objects. The results of PCA come in the form of component scores. By using this formulation the principal components are calculated and the training dataset is reduced. This method used to reduce the training time. The adoption of this method makes the classifier to perform in an efficient time manner. Analyzing the spending behavior pattern of the customer is a promising way to detect the credit card frauds. If any new transactions deviate from this behavior, then those transactions are considered as fraudulent transactions. Each person has a different spending behavior pattern. Behavior based fraud detection model means that the data use in the model are from the transactional behavior of cardholder directly or derived from them. Fraud detection based on the analysis of existing spending behavior of cardholder is a promising way to find the credit card frauds. Based on the spending pattern the customer's usual activities such as transaction amount, billing address etc are learned. Some of the anomalous behaviors include variation of billing address and shipping address, maximum amount of purchase, large transaction done far away from the living place etc.

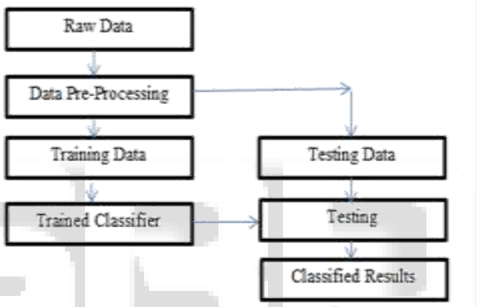


Fig. 3: Behavior based fraud detection using SVM

Initially, the attributes used in the dataset are converted into numerical data. Feature selection is a very important stage in fraud detection. The features in the data efficiently portray the usage behavior of an individual. In this model, the features which interpret the behavior of the customer are selected for detection. Adding irreverent features make the classifier inefficient. Transaction amount is the most important behavior it varies from person to person. Frequency of card usage is calculated from the Date and Time Attributes. Average amount of transactions are calculated from each transactions.

After the completion of the feature selection process, the categorical attributes available in the data are converted to numerical attributes. These numerical attributes are then normalized. The normalized data falls within a defined boundary. This boundary is usually (0, 1) or (-1,+1) for SVM. The appropriate boundary is defined by the SVM designer. After the normalization process, the training and testing data sets are created.

The training set is passed to the SVM. C and γ values are tuned in order to obtain the appropriate pair for the current dataset. The training data is then passed to the SVM for final verification. After obtaining satisfying results, the SVM can be deployed for real-time data processing.

A. SVM Algorithm:

1) Obtain transaction data

```
2)  Preprocess data to convert categorical attributes to numerical attributes
3)  Normalize the numerical data using Min-Max Normalization
4)  Create training and testing data files using the SVM format
5)  Select an attribute a1
6)  Apply K-Means Clustering with respect to the attribute a1
7)  Select the second level attribute a2
8)  For every cluster c obtained
a)  Apply K-Means Clustering with respect to the attribute a2
9)  End for
10) Supply the training file to SVM
11) Set the values for C and γ
12) Obtain results using the current C and γ pair
13) Perform step 6 and 7 till satisfactory results are obtained from the training set
14) Test the accuracy using the test file
15) For every malicious transaction obtained,
a)  Find clusters that contains the current transaction
b)  Using collective animal behavior, check for a similar pattern in the clusters
c)  If the pattern similarity exceeds the threshold t consider the transaction as normal
d)  Else
e)  Consider the transaction as malicious
16) End for
```

IV. Results and Discussion

In the initial detection Model, the unsupervised approach is used. The unknown frauds are easily found by using this approach. The models based on supervised approach must have the labeled data for both normal data and anomalies. It is only able to detect frauds of a type which has previously occurred. In contrast, unsupervised methods don't make use of labeled records. It detects the changes in behavior or unusual Transactions. Unsupervised learning is a feasible method to learn the large and more complex model.

Applying the algorithms in the data, will reduce the number of nearest neighbor searches and number of reach ability distance computation. This model helps detect fraudulent transactions in an efficient manner.

The costs for positive class is always 1 and for the negative class is ratio of negative samples over positive samples are used. Asymmetric error function is used to control the tradeoff between the negative and positive instances. Training samples with different class ratio are used in this work. Increasing the size of the negative classes will not affect the result. Moreover all different samples predict the similar accuracy rate.

	Data samples	**Negative Samples**	**Positive Samples**	**Cost Ratio**
First Set	1000	9	901	90:1
Second Set	1000	46	954	20:1
Third Set	1000	98	902	10:1

Table 1: Data Samples with Different Cost ratio

	Sensitivity	**Specificity**	**g-means**
First Set	0.98	0.8	0.885
Second Set	0.94	0.9	0.919
Third Set	0.95	0.9	0.924

Table 2: Performance Measure

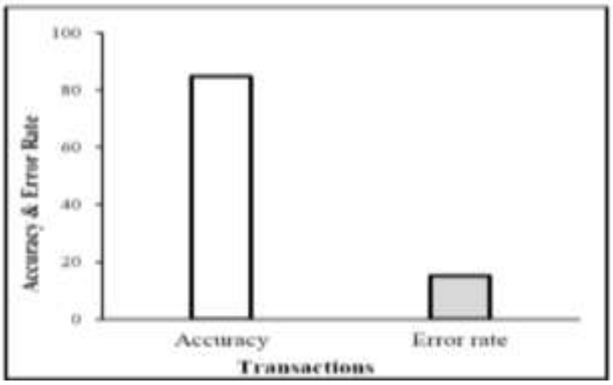


Fig. 4: Accuracy and Error rate

V. Conclusion

The current process provides an efficient way for analyzing the transactions in a bank. This provides an efficient means to detect frauds in a customer's transactions. The currently proposed system has the ability to detect frauds performed by both physical and virtual type of frauds. SVM provides an efficient means for detecting frauds, which is further improvised by incorporating collective animal behavior. The collective animal behavior presents a behavior based analysis to check if users with similar behavior patterns exhibit the same behavior. This helps in further reduction of false positives. Reduction in false positives is one of the main processes that have to be incorporated in any process that involves customers. Analysis of the results shows that our current process provides higher level of accuracy when compared to other fraud detection process.

References

[1] Tareq Allan and Justin Zhan, "Towards Fraud Detection Methodologies", IEEE Proceedings of the Fifth International Conference on Future Technology (Future Tech), 2010.

[2] R. Dhanapal, "An Intelligent Information Retrieval Agent", Elsevier, Knowledge-Based Systems 21, pp. 466-470, 2008.

[3] V. Dheepa, R. Dhanapal and D. Remigious, "A Novel Approach to Credit Card Fraud Detection Model", Journal of Computing, Vol. 2, No. 12, pp. 96, 2010.

[4] Zan Huang, et al., "Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study", Elsevier-Decision Support Systems, Vol. 37, pp. 543-558, 2004.

[5] Kyung-Shik Shin, Taik Soo Lee and Hyun-jung Kim, "An Application Of Support Vector Machines In Bankruptcy Prediction Model", Elsevier, Expert Systems with Applications, Vol. 28, pp. 127–135, 2005.

[6] K. J. Kim, "Financial Time Series Forecasting Using Support Vector Machines", Neurocomputing, Vol. 55(1/2), pp. 307–319, 2003.

[7] Emin Aleskerov, Bernd Freisleben and Bharat Rao, “CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection”, Proceedings of the Computational Intelligence for Financial Engineering, pp. 220-226, 1997.

[8] Zhang yongbin, You Fucheng and Liu Huaqum, “Behavior–Based Credit Card Fraud Detection Model”, Fifth International Joint Conference on INC, IMS and IDC, pp. 855-858, 2009.

[9] Chuang-Cheng Chiu and Chich-Yuan 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.

[10] Wen-Fang Yu and Na Wang, “Research on Credit Card Fraud Detection Model Based on Distance Sum”, Proceedings of the International Joint Conference on Artificial Intelligence, pp. 353-356, 2009.

[11] Tao Guo and Gui-Yang Li, “Neural Data Mining for Credit Card Fraud Detection”, International conference on Machine Learning and Cybernetics, Vol. 7, pp. 3630-3634, 2008.

[12] Suvasini Panigrahi, Amlan Kundu, Shamikr Sural and A.K. Majumadar, “Credit Card Fraud Detection: A Fusion Approach Using Dempster Shafer Theory And Bayesian Learning”, Information Fusion, Vol. 10, No. 4, pp. 354-363, 2009.

[13] Abhinav Srivastava, Amlan Kundu, Shamikr Sural and A.K. Majumadar, “ Credit Card Fraud Detection Using Hidden Markov Model”, IEEE Transactions on Dependable and Secure computing, Vol. 5, No. 1, pp. 37-48, 2008.

[14] Valdimir Zaslavsky and Anna Strizhak , “Credit Card Fraud Detection using Self Organizing Maps”, Information & Security: An International Journal, Vol. 18, pp. 48-63, 2006.

[15] Chen, R., Chiu, M., Huang, Y. and Chen, L, “Detecting Credit Card Fraud By Using Questionnaire-Responded Transaction Model Based On Support Vector Machines”, Proceedings of the Fifth International Conference on Intelligent Data Engineering and Automated Learning, Vol. 3177, pp. 800-806, 2004.

[16] V. Dheepa and R. Dhanapal, “Analysis Of Credit Card Fraud Detection Systems”, International Journal of Recent Trends in Engineering, Vol. 2, No. 3, pp. 126-128, 2009.

[17] Dumais.S, “Using Support Vector Machines For Text Categorization”, IEEE Intelligent Systems, Vol. 13, No. 4, pp. 21–23, 1998.

[18] G. Dror, R. Sorek and S. Shamir, “Accurate Identification Of Alternatively Spliced Exons Using Support Vector Machine”, Bioinformatics, Vol. 21, No. 7, pp. 897–901, 2005.

[19] Osuna. E, “Applying Support Vectors Machines To Face Detection”, IEEE Intelligent Syst. Mag., Support Vector Machines, Vol. 13, No. 4, pp. 23–26, 1998.

[20] Vapnik, V.N, “The Nature of Statistical Learning Theory”, Springer, 1995.

[21] C. Cortes and V. Vapnik, “Support Vector Networks”, Machine Learning, Vol. 20, pp. 1-25, 1995.

[22] Cristianini N and Shawe-Taylor J, “An Introduction to Support Vector Machines and other Kernel-based Learning Methods”, Cambridge University Press, Cambridge, UK, 2000.

[23] Xuchen Li, Lei Wang and Eric Sung, “Ada Boost With Svm–Based Component Classifiers”, Engineering Applications of Artificial Intelligence, Vol. 21, No. 5, pp. 785-795, 2008.

[24] Lindsay I Smith, “A Tutorial on Principal Component Analysis”, Feb 26, 2002.