Xournals

Academic Journal of Forensic Sciences

ISSN: 2581-4273 | Volume 01 | Issue 01 | April-2018

A review of Credit card Fraud Detection techniques in e-commerce

Kaneeka Joshi[1]

Available online at: www.xournals.com

Received 26th December 2017 | Revised 10th February 2018 | Accepted 18th March 2018

Abstract:

With the rise and light growth in e – commerce, the use of credit card for online transactions is also increasing dramatically. Due to this there is a great amount of increase in credit card frauds for which there is a requirement of various detection techniques for determining the fraudulent transactions. Frauds can either be offline or online for regular purchases, credit card is used as a mode of payment. Fraud is considered as the most ethical issue in credit card frauds and it is a million dollar business which is rising every year. Recent advances in techniques based on Data mining, Algorithm system (Genetic algorithm, Artificial Algorithm), Machine learning, Hidden Marokov model are the modern techniques that are introduced for detecting credit card fraudulent transactions. This paper reviews all the fraud detection techniques which have some advantages and disadvantages as well. As according to the study Hidden Markov model and Data mining techniques are considered as the best suitable techniques and may be considered over other techniques successfully.

Key Words: *Credit card, fraud detection, Data mining, Hidden Markov model, and e – commerce,*

Authors:

1. *Sam Higginbottom University of Agriculture, Technology & Sciences, INDIA*

Introduction

Fraud generally refers to obtaining money and goods/services by wrongful or criminal deception or in illegal way which intends to result in personal or financial gain. Fraud deals with criminal events that needs identification which becomes difficult. Due to the development of technology and the increase of internet usage, credit card frauds or can say online frauds are increasing day by day. This wide ranging term "credit card fraud" used for theft or any fraud committed or any transaction that is done as to gain a fraudulent source of funds in a transaction. Credit card fraud is defined as when an individual uses another person's credit card information for his/her personal use without the consent of the owner. These credit card frauds need immediate detection which should be active. In the cases of web technologies the research is necessary as it supports E-commerce in managing and building these applications. This business of E – commerce makes it possible to shop anything anytime as it have no time and geographical restrictions.

Detection of these frauds is difficult task if using normal process of detection, so the use of models i.e., fraud detection models is considered as of more importance in cases of academic or business organizations. E- Commerce are of many kinds – B2B (Business to Business), B2C (Business to Consumer), C2C (Consumer to Consumer) are the three popular application forms of E-commerce. As we have discussed number of models/ systems/ configuration/ process and some preventive measures that are used to avoid credit card fraud which reduces financial risks. These include some common cases such as acquiring or property trading which includes personal and intangible property such as stocks, bonds and copyrights. It is becoming essential to combat these type of fraudulent transactions for which various techniques were applied such as Hidden Markov Model, Artificial Intelligence, Sequence Alignment, Data Mining Techniques, Multiple cryptographic Algorithms and Genetic programming techniques (Rana and Baria, 2015; Meshram and Yenganti).

Types of Frauds

Credit card frauds are mainly divided into two classes:

- *Offline frauds* – are committed by using credit cards that are stolen at place.
- *Online frauds* – are committed via internet, shopping, phone, and web or in case of absence of credit card (Rana and Baria, 2015).

The definition of problems related to credit card frauds

An e-commerce business or system provides suitability for online transactions to shop anything to be a universal business. These frauds can be possible in both ways online as well as offline.

1. While online transactions such as during shopping, fraudster only uses the card information or just want to harass Merchant or to the banks by doing frauds:
2. In case if fraudster don't want to buy anything from the shop but still he/she provides wrong information and payment was done via cash on delivery to harm the merchant.
3. If the credit card of the user was lost or stolen then the fraudster can use the information of card like Credit card number, CVV number to make payment without the consent of cardholder (Rana and Baria, 2015).

Proposed Systems for credit card fraud detection

Various credit card fraud techniques were developed by Researchers.

Data Mining Technique

The technique of Data Mining uses complex data analysis tools that discover the unknown, valid patterns and association among large data sets. The analysis tools include statistical systems, mathematical algorithms and machine learning methods like Neural Networks. This method comprises of collection and maintenance of data which represented in the form of textual, quantitative or in multimedia forms that also involves examination, estimation and prediction. These application uses range of factors to observe data which includes association, classification, sequencing or path analysis. Banks contain large databases which have suitable information on fraudulent acts related to the banking Industry. The important business information was extracted from the stored data and these stored data patterns as clusters naturally inputs some data. The foremost principle while detection is matching the abnormal pattern with the normal pattern. Huge amount of e-commerce industries depends on credit card purchases which is more important for business and organizations (Chaudhary and Mallick, 2012).

Algorithm Technique

Artificial Intelligence (AI), algorithm was used or introduced for the purpose of detecting credit card/ATM cards, cheque books type of fraudulent documents.

Cryptographic Algorithm - According to Meshram and Yenganti, Cryptographic techniques were used for preventing fraud by using tools which asks for secret, questions. The multiple layers of security were enhanced using security questions which wraps the pin number in previous stages using Cryptographic algorithms. With the use of this method fraud can also be prevented (Meshram and Yenganti, 2013).

Genetic Algorithm is another method of detecting credit card frauds as this make use of the existing transactions of credit cards which involves multiple standards for detection. These standards include frequency and location of card usage and balance on credit card book. The filter and priority settings are also responsible for detecting frauds (Alekhya and Basha, 2013).(Bentley et al. 2000) suggested one algorithm that is classified on the basis of credit card transactions into suspicious and non-suspicious classes in genetic programming. Due to the employment of different type of rules were tested in different field in which the best rule is of the high predictability. For the determination of suspect behavior an algorithm was developed in the model cost and rate were evaluated whereas in other cases use of different evaluation techniques are done on the basis of prediction rate and the error rate (Chan et al 1999).

Hidden Markov Model

Hidden Markov Model was used as a detection technique which is used to sequence the model of operation during the credit card transaction which is done by using Baum – Welch algorithm. The transaction was considered as fraudulent if the incoming transaction of credit card was not approved by trained Hidden Markov Model (HMM) with sufficiently large enough probability. Three behaviors of card user were taken into response to detect fraud.

1. Low spending behavior (spending low amount)
2. Medium spending behavior (spending medium amount)
3. High spending behavior (spending high amount)

The spending behaviors are different for different card holder which includes low spending behavior (Ingole and Thool, 2013).The need of training data and test data is always required to detect fraud which uses data mining algorithm K- means internally. Due to this K – means algorithm all credit card information is taken, input data is the number of clusters and these are used in HMM technique. The amount in this technique can be either low, high or medium (Alekhya and Basha, 2013).Abhinav Sruvastava et al proposed a recent research on HMM technique which is a double surrounded random process of probability distribution or pattern with two levels of hierarchy. It can be considered as model which may be more complex process when compared to the traditional Markov model. Many problems exist while using HMM technique such as it does not support many systems or not compatible with some systems.

Outlier Detection

This outlier detection is used for fraud detection system which is generally an observation that diverge from the path of another observations that arises from suspicion. Learning approach that was not supervised is employed by this model which results in a new representation of the observed data. The normal behavior was represented by this method which involves modelling of baseline distribution whereas on other sides i.e. supervised methods in which the models are exercised to differentiate between fraudulent and non-fraudulent transactions through which new observations are assigned. In case of unsupervised methods there is no need to have some prior knowledge about fraudulent and non-fraudulent documents (Chaudhary, Yadav and Mallick, 2012).The outlier detection technique is used where large scale data is involved and this technique works efficiently for applications where computational limitations are present (Pawar et al, 2014). (C. Aggarwal and P. Yu) A method of outlier detection were used for high dimensional data space and the behavior of data of projectiles is required to find the outliers.

Bayesian and Neural network

BNN is a fraud detection technique that is automatic based on machine learning approaches. This techniques does not need to be reprogrammed and this considered as a great advantage of neural network. Neural network processing speed is high in comparison to Bayesian and Neural Network and a high processing time is required for huge neural networks. This fraud detection technique i.e., Bayesian and Neural networks provides good accuracy. The number of disadvantages in case of Neural Network systems such as structure confirmation is difficult, needs excessive training, training efficiency and so on. (Dorronsoro et al 1997) developed a system that is based on the neural classifier system that has limitations that the clustering of data is must done by any type of account. Leonard 1995 explained one disadvantage that is mainly the time constraint while using a rule based expert system in Neural Networks. (Brause et al, 1999) allows the use of neural network technologies which have been used in fraud detection

techniques whereas (Ezawa and Norton, 1996) explained one more technique that is Bayesian Network is also used as an alternative to other methods and applied in telecommunication industries.

Hybridization technique

This is basically sequence alignment of two stages and two analyzers, profile and Deviation analyzer. With the help of Profile analyzer the sequence was matched with the cardholder database whereas in Deviation analyzer, the past history of fraud database was compared with incoming unusual data as the unusual data is passed on to the deviation analyzer. The system was alarmed in case of fraud. The major limitation with this system that it does not help in the detection of cloning cards (Rana and Baria, 2015).Rilly have proposed the technique of BLAST-SSAHA Hybridization technique through which the detection improves by combining both individualities as well as misuse detection.

Literature review

Researchers developed many fraud detection techniques through which the occurrence of frauds are decreasing.

Ghosh and Reilly 1994 advanced neural network for the detection of frauds which should provide a large sample of credit card transaction that is labeled. The cases related to these type are stolen of credit cards or lost cards, counterfeiting, application frauds, non-received issue (NRI) fraud and fraud through mail order.

Alekerovet et al. 1997 recommended in their presentation CARDWATCH that a neural network system based on the data mining process. This is generally a trial product based on data mining process that is developed for credit card fraud techniques and a basic requirement of this system is one network per customer.

Rumelhart 1986 states that organization of nodes into layers and after layering, the layers of neurons were attached with interconnections of modified weight. The new environment of its own behavior matched with the present environment to the new possible situations.

Quinlan, 1993 developed a decision tree method that is a learning system that deals with uniform data and has also developed ID3 method which is more advantageous as it has high flexibility due to which the distribution of data may be done without any assumption and the other one is the robustness which is the main reason for its good utilization.

Fan et al 1999, recommended the use of data mining technique i.e., distributed data mining in fraudulent transactions or detection of credit card frauds. This is considered as the well-organized way of arranging highly distributed databases and Boosting algorithm was used to detect system. Ada Cost is the other name given to Boosting Algorithm which uses classifiers in large amount and more resources of computational devices during detection.

Lane 1999 recommended the use of HMM model to human behavior as after the correct modelling of human behavior, the deviation that is detected must be observed. The behavior of attacker was not found to be similar to authentic user.

Stolfo et al.1999 suggests a Meta learning techniques that can be used for credit card fraud detection systems. This technique of Meta learning combines and divide or integrate the numbers separately to build models. For the detection of fraud and intrusion, the same group worked on it as they use Java agents for this process. The number of important metrics such as the distribution and accuracy of true positive-False positive (TP-FP).To learn models of fraudulent transactions and to acquire findings of high fraud along with low forged alarm this process is considered as suitable for fraud detection.

Syeda et al 2002 proposed a technique that with the use of equivalent large or rough neural network the speed can be enhanced and also suggested the Knowledge discovery process (KDP) that helps in acquire the speed up to 10 processors only but in case of introduction of more number of processors leads to problem of imbalance. For increasing the data mining speed and the discovery process of knowledge use of parallel granular neural networks (PGNNs) is suggested. Due to the complex nature of Hidden Markov Model, the series of time are not scalable or measurable to large size data sets.

According to the suggestion of Cho and Park 2003, HMM based intrusion technique were developed that helps in enhancing the time of modelling and performance. This was improved by the help of privilege transition flows which is based on the knowledge of the domain attacks.

Hoang et al 2003 proposed a new method by using HMM to process sequencing of system calls for anatomical detection.

Joshi and Phoba 2005 examined the abilities of Hidden Markov Model in anatomy detection which classifies the traffic network in the form of attack or normal using HMM.

Abhinav Srivastava et al 2008 suggested that the credit card fraud technique that is Hidden Markov Model (HMM) shows 80% accuracy for acquiring large variations in the input data.

Amalan Kundu et al 2009 suggested that a model of BLAST-SSAHA based on hybridization technique for the online credit card fraud detection. This system improves the detection technique which involves the combination of individualities and misuse techniques of detection.

Alekhya and Basha, 2013 discussed about the E-commerce applications in credit card frauds. The genetic algorithm, a technique applied for the detection of frauds are implemented. In this Java platform was used by a prototypic application and genetic algorithm technique is basically used to detect frauds related to credit cards. This paper preferred Genetic algorithm over other techniques and it is proved by experimental results.

Ingole and Thool, 2013 proposed that HMM technique was used in detection of credit card fraud by sequence modelling of credit card transactions involves the use of the clusters created by clustering algorithm. An HMM technique is used with Baum-Weich algorithm which detects that whether the transaction i.e., incoming is fraudulent or not. The performance of systems was calculated by using metrics and the observed accuracy comes out to be 75%.

Meshram and Yengati 2013 performed experiment on cryptographic algorithm in which a system was proposed that the identity of exact user can be found not only by using security pin number but by embedding secret questions. By selecting the appropriate path for the transferring of any file from source to destination, file can be saved in its own destination and then the transferring process takes place that is fully secured. ‘

Shabbir and Kannadasan 2013 proposed the method of mining technique proves that fraudulent transactions can be deduced and number of false alerts can also be reduced. The genetic algorithm method when applied into the credit card fraud detection systems in bank through which the chances of fraud transactions can be predicted earlier than the credit card transactions.Anti-fraud strategies were applied to avoid frauds related to credit cards.

I. Discussion and Conclusion

(Ingole and Thool, 2013) proved in his research paper that the performance was calculated on the basis of metrics by using HMM technique is 75% whereas according to Abhinav Srivastava et al have proposed 80% accuracy in Hidden Markov Model in input data of large variation . The swindle detection based on outlier detection can detect credit card frauds better than data clustering technique i.e., Hidden Markov Model .Genetic algorithm is considered as a novel technique related to domain of application that when the algorithm was applied onto the bank credit card fraud detections, the probability of fraud transactions may be estimated after the credit card transactions. And through these techniques new strategies or anti-fraud strategies were adopted to prevent from bank frauds and reduce risks (Shabbir and Kannadasan, 2013) but According to Philipi and Sherly,2012 Algorithm techniques has some disadvantages which required specialized and supervised training to sufficiently optimize parameters. With the increase of fraud detection techniques, there is an increase in fraud detection systems.(Pawar et al, 2014)The algorithm techniques has some disadvantages as this requires extensive supervised training, testing and all these requires involvement of human to prepare for training and test cases to sufficiently optimize parameters. Outlier detection techniques were also used which requires large scale data and are used with those applications where memory and computational limitations are present. So (Chaudhery,Yadav and Mallick, 2012) explained in his paper, 13 different classification systems were determined and all these models/system describes the advantages of Data mining technique in Artificial Neural Networks (ANN) for the detection of fraud transactions. As the distribution of training data sets becomes more partial, with this there is a decrease in the performance of all models which directly makes difficult the capturing of fraudulent transactions. (Chaudhey, Yadav and Mallick, 2012) that they considered practice research unusual in case of statistical methods and Neural Network system has advantages and disadvantages as well. The number of disadvantages in case of Neural Network systems such as structure confirmation is difficult, needs excessive training, training efficiency and so on. (Philipi and Sherly, 2012) described in their paper that the algorithm method is considered as more effective in real world as arrangement of data sets is extremely efficient which involves the study of linear and nonlinear relationships that comes directly from modelled data that is in a form of linear fashion. Rama and Baria, 2015 stated in their paper that good accuracy obtained from Bayesian neural networks but required high processing speed and data training to operate it. According to Ghosh and Reilly, 1994, the technique of data mining requires long training time. The proposed limitation for a neural network system based on data mining is that it require one network per customer (Aleskerov et al).Baysian method has advantages of high accuracy, processing speed is good, false alarm reduction, detection rate is improved

and a good application in e- commerce but one disadvantage is that it is highly expensive. According to Rana and Baria, hybridization technique is considered as inexpensive and have high accuracy, HMM and Bayesian neural network technique is quite expensive with medium accuracy. (Rana and Baria, 2015). After reviewing every details about detection technique, we may conclude that Hidden Markov Model, Data mining systems and Genetic Algorithm are the best ways to detect fraudulent transactions.

In this review paper, study is related to the detection of credit card fraud and the techniques related to E-commerce applications. Various detection systems were explored to solve the problem of frauds as the information about several detection methods can help to improve and protecting the applications of E-commerce. Every fraud detection techniques has its own advantages and disadvantages or strengths and weaknesses. In E-commerce, rising of an accurate and credit card fraud technique that is resourceful is necessary. Finally the implementation of some methods such as Genetic algorithm, Hidden Markov Model, Neural Network, Data mining techniques were done. While doing the comparison of performance for predicting the accuracy, this survey or review results that the application used to detect fraud is useful and is used in real world systems. In this study, we characterized the impact of fraud and the techniques of fraud detection methods as well and how these models helps in capturing the fraudulent transactions.

.

References:

Alekhya, P. Phani, and Sk Mahaboob Basha. "*Protecting E-Commerce Systems From Online Fraud*." International Journal of Computer Trends and Technology (IJCTT) – 4.10 (2013): 3549-554. Web. 05 Apr. 2017.

Aleskerov,E., Freisleben, B. & B Rao. 1997., "CARDWATCH: *A Neural Network –Based Database Mining System for Credit Card Fraud Detection*", Proc. Of the IEEE/IAFE on Computational Intelligence for Finance Engineering, 220-226.

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 2009S.

Bentley, P., Kim, J., Jung. G. & J Choi. 2000. Fuzzy Darwinian Detection of Credit Card Fraud, Proc. of 14th Annual Fall Symposium of the Korean Information Processing Society.

Brause R., Langsdorf T. & M Hepp. 1999a. Credit card fraud detection by adaptive neural data mining, Internal Report 7/99 (J. W. Goethe-University, Computer Science Department, Frankfurt, Germany).

Chaudhary , Khyati, Jyoti Yadav, and Bhawna Mallick. "*A review of Fraud Detection Techniques: Credit Card* ." International Journal of Computer Applications 45.1 (2012): 39-44. Web. 05 Apr. 2017.

Chaudhary, Khyati, and Bhawna Mallick. "*Credit Card Fraud: Bang in E-Commerce*." International Journal of Computational Engineering Research 2.3 (2012): 935-41. Web. 05 Apr. 2017.

Chaudhary, Khyati, and Bhawna Mallick. "*Credit Card Fraud: The study of its Credit Card Fraud: The study of its Fraud: The study of its impact and detection techniques* ." International Journal of Computer Science and Network 1.4 (2012): 31-35. Web. 05 Apr. 2017.

Ingole , Avinash , and R. C. Thool. "*Credit Card Fraud Detection Using Hidden Markov Model and Its Performance.*" International Journal of Advanced Research in Computer Science and Software Engineering 3.6 (2013): 626-32. Web. 05 Apr. 2017.

Meshram , Pratiksha L., and Traun Yenganti. "*Credit and ATM Card Fraud Prevention Using Multiple Cryptographic Algorithm.*" International Journal of Advanced Research in Computer Science and Software Engineering 3.8 (2013): 1300-305. Web. 05 Apr. 2017.

Nimisha Philip, Sherly K.K, "Credit Card Fraud Detection Based on Behaviour Mining" TIST.Int.J.Sci.Tech.Res., Vol.1 , 2012, pp. 7- 12.

Philip, Nimisha, and Sherly K. K. "*Credit Card Fraud Detection Based on behavior mining.*" TIST.Int.J.Sci.Tech.Res 1 (2012): 7-12. Web. 05 Apr. 2017.

QUINLAN, J. R. (1993): C4.5: Program for machine learning. Morgan Kaufmann, San Mateo, CA, USA.

Rana , Priya J., and Jwalant Baria. "*A Survey on Fraud Detection Techniques in Ecommerce*." International Journal of Computer Applications 113.14 (2015): 5-7. Web. 05 Apr. 2017.

Reddy, P. Amarnath , and K. Srinivas. "*Credit Card Fraud Detection and Alerting Using Hidden Mark Over Model And Sms Gateway*." International Journal of Engineering Research & Technology 1.8 (2012): 1-7. Web. 05 Apr. 2017

S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int'l Conf. System Sciences:Information Systems: Decision Support and KnowledgeBased Systems,vol. 3, pp. 621-630, 1994.

S. Stolfo and A.L. Prodromidis, "Agent-Based Distributed Learning Applied to Fraud Detection," Technical Report CUCS-014-99, Columbia Univ., 1999.

S.S. Joshi and V.V. Phoha, "Investigating Hidden Markov Models Capabilities in Anomaly Detection," Proc. 43rd ACM Ann. Southeast Regional Conf., vol. 1, pp. 98-103, 2005.

Srivastava, Abhinav, Kundu, Amlan, Sural, Shamik and Majumdar, Arun K., (2008) "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable and Secure Computing, Vol. 5, No. 1, pp. 37-48.

Syed, Shabbir Ahsan, and R. Kannadasan. "*An Effective Fraud Detection System Using Mining Technique.*" An Effective Fraud Detection System Using Mining Technique 3.5 (International Journal of Scientific and Research Publications): 1-4. Web. 05 Apr. 2017.

Syeda, M., Zhang, Y. Q., and Pan, Y., 2002 Parallel Granular Networks for Fast Credit Card Fraud Detection, Proceedings of IEEE International Conference on Fuzzy Systems, pp. 572- 577 (2002).

T. Lane, "Hidden Markov Models for Human/Computer Interface Modeling," Proc. Int'l Joint Conf. Artificial Intelligence, Workshop Learning about Users, pp. 35-44, 1999.

W. Fan, A.L. Prodromidis, and S.J. Stolfo, "Distributed Data Mining in Credit Card Fraud Detection," IEEE Intelligent Systems, vol. 14, no. 6, pp. 67-74, 1999.

X.D. Hoang, J. Hu, and P. Bertok, "A Multi-Layer Model for Anomaly Intrusion Detection Using Program Sequences of System Calls," Proc. 11th IEEE Int'l Conf. Networks, pp. 531-536, 2003.

THIS PAGE IS
INTENTIONALLY
LEFT BLANK