

"Credit Card Fraud Detection Algorithm: Survey"

Ms. DurgaWanjari, Prof. Rajesh Nasare

(Computer Science & engineering, Rajiv Gandhi College of Engineering, Nagpur.

Email: durgawanjari2012@gmail.com)

(Computer Science & engineering, Rajiv Gandhi College of Engineering, Nagpur

Email : rajeshnasare@gmail.com)

Abstract:

Credit card is the basic requirement for the today's world. Everything is connected to internet and there is a basic need for the online shopping. As fraudulent are increasingly day by day. If someone stolen user's credentials, then he/she will able to access personal data of user. So, in this paper we proposed a system based on credit card frauds. This system used Hidden markov model and Genetic algorithm to finding credit card frauds in system. HMM is used for spreading patterns of transactions and concurrently report the time stamp and IP address of the intruder machine and Genetic algorithm is used for providing better solution of the given problem.

Keywords —Fraud Detection, Genetic Algorithm, Hidden markov model, online shopping, spreading patterns.

I. INTRODUCTION

Online activities are well acknowledged to every citizen of the society with the eminent growth of e-commerce. Online activities mainly involve regular purchase of goods, electronic devices and other such things. The online transactions made for such activities are secure payment methods that authorize the transfer of funds.

Apart from its magnificent advantage they do face some of their pitfalls regarding the security. The illicit use of these credit cards is a major issue to ponder on. The credit card fraud can be done for various reasons, mainly to get unaccredited fraud from the account.

It is thus the responsibility of the bank to safeguard the amount transferred online on the internet of the card holder. The bank organization can adopt various existing methodologies such as case based reasoning for fraud detection in order to reduce the financial loss.

These transactions are supported by different bank cards which makes the operation easy. A huge population use credit card for its undemanding accessibility. The bank has accumulated a vast count of credit card transactions.

II. LITERATURE SURVY

Aayushi Gupta, Dhananjay Kumar, AtulBarve has proposed the HMM, is use the normal behavior of a cardholder needs an enough number of normal transactions and fraud transactions for learning fraud patterns. To make it more effective we have enclosed the provision of determining the IP address of intruder machine along with its time stamp. HMM uses cardholders spending behavior to detect fraud. High spending behavior, Medium spending behavior, Low spending behavior. Different users have their various spending behavior (high, medium, low). High spending behavior of any user shows that cardholder spending amount(H), medium spending behavior of any user shows that user spend medium amount(M),

low spending behavior of any user shows that cardholder spend low amount(L).

RuchiQberoi has proposed the problem and also generates the result of the fraudulent transactions. Genetic algorithm is used to make decision about the network topology, number of hidden layers, and number of nodes that can be used in designing the neural network for solving our credit card fraud detection problem. The Genetic algorithms are evolutionary algorithms in which the aim is to obtain the better and optimal solutions. In the fraud transactions can be predicted soon after credit card transactions by the banks.

R. Chantal and P. Gayathri has proposed Decision Tree algorithm which is data mining induction technique that recursively divide a dataset of records using depth first greedy approach. A decision tree structure is made of root, internal nodes, leaf. The tree structure is used in sorting unknown data records. In this method, a credit card fraud detection using useful algorithm for Decision Tree Learning and focus is on the information growth based. This method estimates the best split of purity measures, entropy and information gain ratio to test the best classifier attribute. The author simply find out the fraudulent user through tracing fake mail and IP address. Customers are suspicious if the mail is fake and they traced all details about the sender through IP Address.

RaghavendraPatidar and L. Sharma has proposed fraud detection using Neural Network is totally based on the human brain working principle. There is a fix pattern of credit card use, made by the way consumer uses a credit card. The Neural Network based fraud detection system check for the pattern used by the fraudster and matches with the pattern of the valid card holder on which Neural network has been trained. If the pattern matches, then the Neural Network declare the authorize transaction.

Joseph Pun. YoriLawryshyn has proposed the problem and limitations of the original method. Themeta-learning solution that improves Bayesian

Network at both high-level and system-level views. In addition to discussing the methodology of the proposed solution, this essay studies other attributes of meta-learning implementations, from the advantages and challenges to key design considerations and performance comparison to state-of-the-art solutions in the market.

III. CREDIT CARD FRAUD DETECTION METHODS

This method is used to find out whether the given transaction is agenuine or fraudulent. In case if the transaction is fraudulent, then the detection system must recognize the fraudulent transaction and produce an alert /alarm for such transactions. There are various methods that can be used to detect credit card fraud detection. Some of these are as follows:

1. Hidden markov model
2. Genetic Algorithm
3. Decision Tree
4. Neural Network
5. K-means clustering algorithm
6. Meta Learning Strategy

1.Hidden Markov Model (HMM): A Hidden Markov Model is a finite set of states; every state is associated with a probability distribution. Transitions among these states are administered by a set of probabilities called transition probability. In a specific state a possible out come or observation can be produced which is associated symbol of observation of probability distribution.

1. Low spending behavior
2. Medium spending behavior
3. High spending behavior

HMM produces high false alarm as well as high false positive.

This is the mathematical formula of hidden markov model.

Hidden states $Q = \{ q_i \}, i = 1, \dots, N.$

Transition probabilities $A = \{ a_{ij} = P(q_j \text{ at } t+1 \mid q_i \text{ at } t) \},$ where $P(a \mid b)$ is the conditional probability of

agiven $b, t = 1, \dots, T$ is time, and q_i in Q .

Informally, A is the probability that the next state is q_j given that the current state is q_i .

Observations (symbols) $O = \{ o_k \}, k = 1, \dots, M$.

Emission probabilities $B = \{ b_{ik} = b_i(o_k) = P(o_k | q_i) \}$, where o_k in O . Informally, B is the probability that the output is o_k given that the current state is q_i .

Initial state probabilities $\Pi = \{ p_i = P(q_i \text{ at } t = 1) \}$.

Algorithm steps:

Training phase: cluster creation

STEP 1: To identify the profile of cardholder from their purchasing.

STEP 2: The probability calculation depends on the amount of time that has elapsed since entry into t4 current state.

STEP 3: To Construct the training sequence for training model.

Detection phase : Fraud detection

STEP 1: To generate the observation symbol.

STEP 2: To form new sequence by adding in existing sequence.

STEP 3: To calculate the probability difference and test the result with training phase.

STEP 4: Finally ,If both are same it will be a normal customer else there will be fraud signal will be provided.

2. Genetic Algorithm: The Genetic algorithms are evolutionary algorithms which aim to obtain the better solutions to technically eliminate the fraud, a high importance have been given to develop secure and efficient –payment system to detect whether a transaction is fraudulent or not. Genetic algorithm used in data mining mainly for variable selection and mostly coupled with other data mining algorithms. And their combination with other techniques has a very good performance. Genetic algorithm has been used in credit card fraud detection number of transactions and is easy accessible for computer programming language implementation, thus make it strong in credit card fraud detection. But this method has high performance and is quite expensive.

This is the mathematical formula of genetic algorithm.

Selection of population through the transaction.

Mutation probability (mp):

Number of unique transaction / Number of total transaction in a particular year.

Fitness formula (fy):

Sum of all the unique transaction present for mutation / number of unique transaction.

Algorithm:

STEP 1: Input group of data credit card transactions, every transaction record with n attribute, and standardize the data , get the sample finally ,which includes the confidential information about the cardholder.

STEP 2: Compute the critical values, calculate, the credit card usages frequency count. Credit card overdraft, current bank balance, credit card usage location, average daily spending.

STEP 3: Generate critical values fraud after limited number of generations. Critical fraud detected, monitor able fraud detected, ordinary fraud detected etc. Using genetic algorithm.

STEP 4: Generate fraud transactions using this algorithm this is to analyze the feasibility of credit card fraud detection based on techniques, then applies detection mining based on critical values into credit card fraud detection and proposes this detection produces and its process.

3. Decision Tree: A decision tree structure is used in classifying unknown data records. In deciding out the fraudulent customer /merchants through tracing fake mail and IP address. These trees can be planted via machine learning based algorithms such as the ID3, and C4.5 and MLPC which are applied on credit card database. The core of decision tree model is to construct a decision tree is to construct a decision tree with high accuracy and small scale. If the new transaction is the same type of fraud, then the node become a leaf and is labelled as fraud. This model is very fast and has a high flexibility.

4. Neural Network: Fraud detection using neural network is totally based on the human brain working principal. In neural network is a fix pattern of credit card use, made by the way customer uses it's credit card. System check for the pattern used by the fraudsters and matches with the pattern of the original cardholder on which the neural network, if the pattern matches the neural network declare the authorize transactions. It extract rules and predict future activity based on the current situation. Neural network, effectively, banks can detect fraudulent use of a card, faster and more efficiently.

5. K-means clustering: *k- means* Clustering is a process of arrangement data in to group of similar objects. Different grouping results are obtained from various clustering methods available to group the dataset. The choice of a particular method will depend an the desired output.

The clustering methods are:-

1. Hierarchical Agglomerative methods
2. Partitioning methods
3. The single link method(SLINK)
4. The complete link method(CLINK)
5. The group average method.

6. Meta- Learning strategy: The objective of meta learning is to generate a user supporting system for selection of the most appropriate supervised learning algorithms for tasks. It can be filter the legitimate transactions from the fraudulent ones, and by quickly and accurately identifying the fraudulent transactions, fraud losses can be reduced." Meta-Learning" techniques introduced by chan and stalfo. In the combiner strategy the attributes and correct classifications of credit card transaction instances are used to train multiple base classifiers.

IV. DIFFERENT TYPES OF ALGORITHM / TECHNIQUES FOR CREDIT CARD FRAUD DETECTION SYSTEM.

Author	Year	Techn	Description
--------	------	-------	-------------

Name	Year	Techniques	Description
Aayushi Gupta,Dh ananjay Kumar, AtulBarve	2017	Hidden markov model	Hidden markov model is use the normal behavior of a card holder needs on enough number of normal transaction and fraud transactions for learning fraud patterns.
RuchiQb eroi	2017	Genetic Algorithm	GA use for the providing better solution of the given problem. And also generates the results of the fraudulent transaction. GA is used to make decision about the network topology, number of hidden layer, and number of nodes that can be used in designing the neural network for solving our credit card detection system.
R.Dhanpal and P. Gayatri	2012	Decision tree/Hunt algorithm	Decision tree algorithm which is data mining induction technique that recursively divide a dataset of records using depth first greedy approach. A

			decision tree structure is made of root internal nodes leaf. Author simply find out the fraudulent user through tracing fake mail and ip address.
R.Patidar and L.Sharma	2011	Neural Network/Back propogation algorithm	Neural network is totally based on the human brain working principle there is a fix pattern of credit card use, made by the way consumer user a credit card. Neural network based fraud detection system check for the pattern used by fraudster and master with the pattern of the valid card holder on which neural network has been trained.
Vaishali	2014	K-means clustering algorithm	K-means clustering algorithm is an unsupervised technique. Unsupervised technique is useful when there is no perior knowledge about the particular class of observations in a data set. K-means clustering is a simple and

			efficient method to cluster of data.
Joseph Pun. YoriLawr yshyn	2012	Meta learning strategy	Improvement in catch fraud than neural network

V. COMPARSION TABLE FOR CREDIT CARD FRAUD DETECTION ALGORITHM/TECHNIQUES

Algorithm	Speed of detection	Accuracy	cost
Hidden markov model	Fast	Low	Expensive
Genetic Algorithm	Good	Medium	Inexpensive
Decision Tree	Fast	Medium	Expensive
Neural network	Fast	Medium	Expensive
K-means clustering	Fast	Fast	Expensive
Meta learning strategy	Fast	High	Expansive

VI. CONCLUSION

The technology supporting online transaction has provoked the use of payment cards. There are various fraud detection techniques. The Bank database of user transaction can be used for study the user transaction pattern to avoid the fraud transaction. Hidden markov model and Genetic algorithm is better than all other algorithm discussed here.

REFERENCES

[1] Ayushi Gupta and Dhananjay Kumar and AtulBarve , "Hidden Markov Model based credit card fraud detection system with time

- stamp and ip address" , International journal of computer applications (0975-8887) Volume 166-No.5 , May 2017.
- [2] Ruchi Oberoi, "credit -card fraud detection system: using Genetic algorithm "International journal of computer and Mathematical Science IJCMS ISSN 2347-8527 Volume 6, Issue 6 June 2017.
- [3] R. Dhanpal and P. Gayatri, "Credit Card Fraud detection using decision tree for tracing email and ip,"International Journal of Computer Science Issues, Vol. 9, no. 2, 2012.
- [4] R. Patidar and L. Sharma, "Credit Card Fraud Detection Using Neural Network", International Journal Of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issus-NCIAI2011, June 2011.
- [5] Vaishali "Fraud Detection in Credit Card by Clustering Approach", International Journal Of Computer Applications (0975-8887) Volume 98-No.3, July2014.
- [6] N. Malini and M. Pushpa "Analysis on Credit Card Fraud Identification baesd on KNN and outlier detection" 3rd International conference on advance in electrical, electronics, information, and communication and Bio-informatics (AEEEEICB17).
- [7] Shraddha Asati, Rajesh Nasare, and Hemant Turkar "Image forgery detection based on illumination inconsistencies" International Journal of Innovative Research in computer and communication Engineering, vol-5, Issue 4, April 2017.
- [8] Rajesh Nasare "Spam Mail Detection Using Artificial Neural Network" Imperial journal of Interdisciplinary research, vol-2, Issue-5, 2016.
- [9] Josef Pun, Yuri Lawryshyn "Improving Credit card fraud detection using a meta-classification strategy", International Journal of computer applications (0975-8887) volume 56-No.10, October 2012.
- [10] I. Srivastava and A. Kundu, "Credit Card Fraud Detection using hidden markov model", IEEE Transaction on Dependable and Secure Computing, Vol. 5, no. 1, 2008.
- [11] Ghosh S., Reilly D. L., "Credit Card Fraud Detection with a Neural -Network" Proceedings of the International Conference on System Science, pp.621-630,1994.
- [12] Aleskerov E., Freisleben B., and Rao B., "CREDITWATCH: A Neural Network Baesd Database Mining System For Credit Card Fraud Detection", Proc.
- [13] Dorronsor J.R., Francisco G., Carmen S., and Carlos S.C., "Neural Fraud Detection in CeditCard Operation" IEEE Transaction on Neural Network, VOL.-08, NO.-4, and pp.: 827-834, 1997.
- [14] Kokkinaki, A., "On Atypical database Transactions: Identification of Probable Fraud using Machine Learning for User Profiling." Knowledge and Data Engineering Exchange Workshop. IEEE, PP.:107-113, 1997.
- [15] Stolfo S.J., Fan D.W., Lee W., Prodromidis A.L., and Chan P.K., "Credit Card FraudDetection Using Meta-Learning: Issues andInitial Results", Proc. AAI Workshop AI Methods in Fraud and Risk Management, pp.:83-90, 1997.
- [16] Brause R., Langsdorf T., and Hepp M., "Neural Data Mining for Credit Card Fraud Detection", Proc. IEEE Int'l Conf. Toolswith Artificial Intelligence, pp.:103-106, 1999.
- [17] Kim, M, and Kim T., "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection", Proceedings of IDEAL. pp.:378-383, 2002.
- [18] Chiu A., Tsai C., "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.
- [19] Foster and Stine R., "Variable Selection in Data Mining: Building a Predictive Model for Bankruptcy", Journal of American Statistical Association, pp.: 303-313, 2004.
- [20] K.RamaKalyani, D.UmaDevi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm", International Journal of Scientific & Engineering Research Volume 3, Issue 7, July-2012.