International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Card Fraud Detection System in Payment Gateway by HMM

Dhanashree Devendra Surve and Prof. Chaitanya Mankar

Dhole Patil College Of Engineering Pune ,India.

Received 10 Nov 2020, Accepted 10 Dec 2020, Available online 01 Feb 2021, **Special Issue-8 (Feb 2021)**

Abstract

The evolution of the new technology supports the online transactions to be held with the assistance of various payment cards. Credit card frauds have become increasingly constant in living years and critical for banks to enhance fraud detection so as to protect their cardholders from financial loss. The simple way to detect such kind of fraud is to decipher the spending pattern on each card and to highlight any irregularity with respect to the "standard" spending pattern. In this paper we try to review Hidden Markov model which works on such technique. The HMM, trained with 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. The simulation analysis include different real dataset to identify the fraud and discover the intruder. Form model it is proven that it works with more efficiency than existing models.

Keywords: *Hidden Markov model, spending pattern, fraud transaction, credit card, time stamp, financial loss.*

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. 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.

Apart from its magnificent advantages 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 funds 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, decision tree, and neural network for fraud detection in order to reduce the financial loss [1].

Literature Survey

Credit card fraud detection has been a current evoking issue of major concern. In affect to this various detection techniques such as genetic algorithms, data mining, neural networks, clustering techniques and decision tree are used.

Ghosh and Reily [3] implemented the neural network system which involved cases dealing with lost cards, stolen cards, stolen card details, application fraud etc. Aleskerov, Freisleben and Rao [4] also developed a system on neural network called Card watch. The system focus towards commercial implementation. Dorronsoro and others [5] developed a neural network based detection system called Minerva. This system proposes the facility to ingrain itself deep in credit card transaction servers to detect fraud in real-time. Kokkinaki [6] suggested to create a user profile for each credit card account and to test incoming transactions against the corresponding user's profile. Chan and Stolfo [7] studied the class distribution of a training set and its effects on the performance of multi classifiers on the credit card fraud domain. Brause and others [8] looked specifically at credit card payment fraud and identified fraud cases by combining a rule-based classification approach with a neural network algorithm. Kim [9] proposed a fraud density map technique to improve the learning efficiency of a neural network. Chiu and Tsai [10] identified the problem of credit card transaction data having a natural skewness towards legitimate transaction. Foster and Stine [11] attempted to predict personal bankruptcy using a fully automated stepwise regression model.2010 Accounts

Proposed Methodology

The Hidden Markov model is undemanding and easily manageable sequential model which is use to model

the spending convention of the card holder (user). It is a doubly embedded random process comprising of two disparate levels. One of them remains hidden and other is noticeable to observer. The Hidden Markov model has greater potential in managing complex process than the traditional Markov model. The considerable advantage seen in the model is the diminution in number of FP (False Positives). FP is the transition identified as fraudulent by the fraud detection system but although they are genuine. The new model consists of finite set of states which are associated with probability distribution. Transitions among different states are supervised by set of probability called as transition probability [13]. Every state in model originates some outcome called as observation calculated according to corresponding probability distribution. HMM can be successfully applied to various applications in temporal pattern recognitions such as speech, handwriting gesture reorganization part of speech tagging and bioinformatics [12].

A. The HMM can be well defined with the following elements-

• N number of states that are hidden denoted by a set S = {S1,S2,S3,…. SN}, where i =1, 2,……, N, is count of state and Si, is an individual state.

• M denotes the total number of observation symbols. When observations are continuous then M is infinite. We denote the set of symbols V = $\{v_1,v_2,\ldots v_M\}$ where v_i, is an individual symbol.

• A set containing probability of moving from one state to another, defined as transition probability.

aij= P{qt+1 = Sj | qt = Si}, 1≤i, j≤N where q_t denotes the present state.

aij≥ 0, 1≤i, j ≤ N

and = 1, 1≤ i ≤ N

• Matrix B, indicating observation symbol probability A probability distribution in each of the states is given as, bj(k) = P{at =Vk| qt = Sj},1 ≤ j ≤ N, 1 ≤ k ≤ M where, Vk denotes the k_{th} observation symbol and a_t the current parameter vector. The given equation should satisfy some constraints bj(k) ≥ 0, 1 ≤ j ≤ N, 1 ≤ k ≤ M and

j(k) =1, 1 ≤ j ≤ N

• The initial state probability given by Π = {Πi} where,

Πi = P {qi = Si}, 1 ≤ i ≤ N i= 1

After knowing all these elements, the HMM is ready to work. We consider the initial sequence of transaction of card holder. The transactions made by the credit card holder is categorized into three clusters l, m, h for low medium and high category transaction respectively. The volume of each cluster is resolved considering the limit of the credit card. The amount up to 35% of card limit belongs to l cluster, upto 65% belongs to m and above that comes under h cluster.

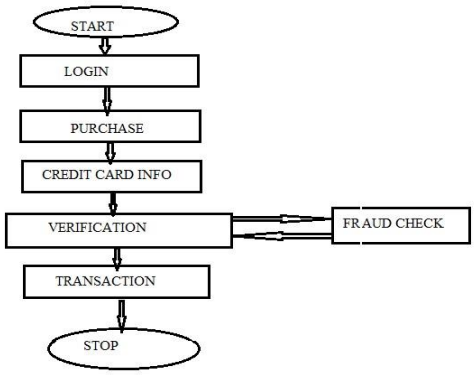

Fig. System Architecture

Let O1,O2,……..OR be consisting of R symbols to form a sequence. The HMM now works in the following manner

• To calculate the probability of acceptance (alpha1) we consider a series of last R transaction of the card holder. The value is given by α1=P (O1, O2, O3,……OR|λ)

• Let OR+1 is the new generated transaction at time t+1. The total no of sequences R+1. Now for calculating the second probability of acceptance we will drop O1 observation and the sequence will be from O2 to OR+1 α2= P (O2,O3,O4,……..OR+1|λ)

• Next we find the standard deviation ▲α=α1-α2

• The standard deviation calculated is now compared with threshold value. If percentage change in probability is found more than predefine threshold value then the transaction will be consider as fraud transaction.

If ▲α/α1≤ threshold value (Θ)

• Ideally the threshold value is taken as 0.5 and further the value modifies every time the algorithm runs using the below formula

Threshold=(▲α/α1+threshold)/2

Due to the obvious security reasons it is very difficult to fetch the dataset from any bank. So in order to get our results simulated analysis is performed by considering a random dataset of transactions for any credit card holder. Firstly, all the transaction sequence

need to be categorized into three clusters namely low medium and high according to the user credit card limit. Assuming the credit card limit to be ₹. 10000 in our case, the range of clusters thus produced will be low {₹. 0, 3000}, medium {₹. 3000, 6000} and high {₹. 6000, 10000}. After deciding the categories the fraud detection of incoming transaction will be verified by last 10 transactions. The system has the flexibility for the future enhancement at the same time shows its advantage of dynamic nature. There will always be a method to enhance the probability which we use for the fraud detection based on practical datasets and values. Also the algorithm used is applied at one layer. For stronger protection multiple layer algorithm can be implemented. Further in future we can design an application which can be add sophisticated modules like capturing the photo of the attacker.

Comparative studies revealed an accuracy of the system to be about 80% for a wide range of input dataset. The percentage change in the probabilities of previous and new transaction sequence is compared with the threshold value which decides whether the upcoming transaction is fraudulent or not. Thus the produced system is genuine to a great extent. It has also reduced the complexity when compared with the existing system. In our simulation analysis we have considered a small set of data, but our proposed system is capable of handling larger range of transactions which is quite certain in real life scenarios.

Result and Discussions

In our proposed model, we have found out more than 84% transactions are genuine and very low false alarm which is about 7 % of total number of transactions. The relative studies and our results sure that the correctness and effectiveness of the proposed system is secure to 80 percent over a broad deviation in the input data. At the initial state HMM checks the upcoming transaction is fraudulent or not and it allow to accept the next transaction or not based on the probability result. The different ranges of transaction amount like low group, medium group, and high group as the observation symbols were considered. The types of item have been considered to be states of the Hidden Markov Model. It is recommended that a technique for finding the spending behavioral habit of cardholders, also the application of this knowledge in deciding the value of observation symbols and initial estimation of the model parameters

Conclusions

In this paper, it has been discussed that how Hidden Markov Model will facilitate to stop fraudulent online transaction through credit card. The Fraud Detection System is also scalable for handling vast volumes of transactions processing. The HMMbased credit card fraud detection system is not taking long time and having complex process to perform fraud check like the existing system and it gives better and fast result than existing system. The Hidden Markov Model makes the processing of detection very easy and tries to remove the complexity. At the initial state HMM checks the upcoming transaction is fraudulent or not and it allow to accept the next transaction or not based on the probability result. The different ranges of transaction amount like low group, medium group, and high group as the observation symbols were considered. The types of item have been considered to be states of the Hidden Markov Model. It is recommended that a technique for finding the spending behavioral habit of cardholders, also the application of this knowledge in deciding the value of observation symbols and initial estimation of the model parameters

References

[1]. Khyati Chaudhary, Bhawna Mallick, "Credit Card Fraud: The study of its impact and detectionTechniques", International Journal of Computer Science and Network (IJCSN), pp: 3135, 2012.

[2]. Bilonikar Priya, "Survey on Credit Card Fraud Detection Using Hidden Markov Model", International Journal of Advanced Research in Computer and Communication Engineering 2014.

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

[4]. Ghosh, S., and Reilly, D.L., 1994. Credit Card Fraud Detection with a Neural-Network, 27th Hawaii International Conference on Information Systems, vol. 3 (2003), pp. 621-630.

[5]. 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).

[6]. Stolfo, S. J., Fan, D. W., Lee, W., Prodromidis, A., and Chan, P. K., 2000. Cost-Based Modeling for Fraud and Intrusion Detection: Results from the JAM Project, Proceedings of DARPA Information Survivability Conference and Exposition, vol. 2 (2000), pp. 130-144.

[7]. Aleskerov, E., Freisleben, B., and Rao, B., 1997. CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection, Proceedings of IEEE/IAFE: Computational Intelligence for Financial Eng. (1997), pp. 220-226.

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

[9]. Brause, R., Langsdorf, T., and Hepp, M., 1999. Neural Data Mining for Credit Card Fraud Detection, Proceedings of IEEE International Conference Tools with Artificial Intelligence (1999), pp. 103-106.

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

[11]. Phua, C., Lee, V., Smith, K., and Gayler, R., 2007. A Comprehensive Survey of Data Mining-Based Fraud Detection Research (2007), March.

[12]. Rabiner, L.R. 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of IEEE, vol. 77, no. 2 (1989), pp.257-286.

[13]. Ourston, D., Matzner, S., Stump, W., and Hopkins, B., 2003. Applications of Hidden Markov Models to Detecting Multi- Stage Network Attacks, Proceedings of 36th Annual Hawaii International Conference System Sciences, vol. 9 (2003), pp. 334-344.

[14]. Cho, S.B., and Park, H.J., 2003. Efficient Anomaly Detection by Modeling Privilege Flows Using Hidden Markov Model, Computer and Security, vol. 22, no. 1 (2003), pp. 45-55.

[15]. Kim, M.J., and Kim, T.S., 2002. A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection, Proceedings of International Conference on Intelligent Data Eng. and Automated Learning, (2002), pp. 378-383.

[16]. Kaufman, L., and Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis, Wiley Series in Probability and Math. Statistics, (1990).