www.ijecs.in
International Journal Of Engineering And Computer Science
Volume 8 Issue 11 November 2019, Page No. 24878-24882
ISSN: 2319-7242 DOI: 10.18535/ijecs/v8i10.4386

Credit Card Fraud Detection Using Hidden Markov Model

Mandeep Singh, Sunny Kumar, Tushant Garg

Assistant Professor CSE Department
HMR Institute of Technology & Management, GGSIPU New Delhi
Student CSE Department
HMR Institute of Technology & Management, GGSIPU New Delhi
Student CSE Department
HMR Institute of Technology & Management, GGSIPU New Delhi

Abstract
Now a day the usage of credit cards and net banking for online payments has dramatically increased. The most popular mode of online as well as regular purchase payments is through credit card and security of such transactions is also a major issue as frauds are increasing rapidly. In the existing scenario, fraud is detected after the transaction is done and it makes more difficult to find out fraudulent loses barred by issuing authority. In this paper, we observe the behavior of credit card transactions using a Hidden Markov Model (HMM) and show how it detects frauds. An HMM is initially trained with the normal behavior of transactions. If the present credit card transaction is not accepted by the trained HMM with enough high probability, then it declares as a fraudulent transaction. At the same time, we try to ensure that no genuine transactions are rejected.

Keywords: *Credit card frauds, HMM, Online fraud detection, Electronic commerce.*

1. Introduction

Credit-card-based purchases can be divided into two types: offline and online transactions. The cardholder needs his card physically for a merchant to make a successful transaction in an offline purchase. An attacker must steal the credit card for fraudulent transactions in an offline purchase. In the second kind of purchase, only some important information about a card (card number, expiration date, card verification value) is required to make a successful payment. To commit online fraud, a fraudster simply needs the card details. Sometimes, the original cardholder is not aware that someone else has stolen or seen his card information. The effective way to detect these types of frauds is to analyze the spending behavior of every card and to figure out any inconsistency with respect to normal spending behavior. Detection of fraud based on the analysis of existing purchase data of cardholders is an effective way to reduce the rate of successful card frauds. Since humans tend to exhibit specific behaviorist profiles, every card-holder can be represented by a set of patterns containing information about the time since the last purchase, the typical purchase category, the amount of money spent, etc. Deviation from such patterns leads to a potential threat to the system. According to a [1] survey by Google, Pune suffered the most from credit card frauds in India with Mumbai on second. Hackers take the benefit of the cashless economy to achieve the frauds. Some ways like skimming, phishing, scanning, etc. by which money can be taken off from your credit card. Due to a lack of awareness, people share their card details with fraudulent emails and make easier to happen fraud. Here it is shown that how HMM be used to detect fraud and create a secure online transaction as HMM trained with the normal behavior of card activities and uses that past behavior to detect fraudulent transactions.

2. Literature Review

Detection of credit card fraud has attracted a lot of attention in researching, and several methods have been proposed, with particular emphasis on neural networks, data mining, and distributed data mining.

Ghosh and Reilly[8] have proposed a fraud detection technique with a neural network for all kinds of credit or debit cards. It was qualified on a large scale sample of labeled transaction. All these kinds of transactions were authorized due to card loss, stolen Cards, application fraud, mail-order fraud, Non-receive issue (NRI) fraud. Syeda et al.[9] have used Parallel granular neural networks (PGNNs) to boost data mining rates. To this effect, a total framework has been enforced. Stolfo et al.[10] suggested credit card fraud detection (FDS) using meta-learning techniques that help to learn different models of fraudulent transactions. Meta-learning is a scheme that offers a way for many separately construct classifiers to be combined and incorporated. A meta classifier will be equipped to compare the base classifier predictions. Another same group also worked on a fraud detection cost-based model. Meta-learning java agents were used, which is a distributed data mining platform for the detection of credit card fraud. They identified some type of important performance matrices as True-positive-False Positive (TP-FP) spread and accuracy.

Aleskerov et al.[11] present CARD WATCH, a database mining program used to detect credit card fraud. Based on a neural learning module, the system provides a number of commercial databases with an interface. Kim and Kim[12] identified legitimate and fraudulent transactions as the two main reasons for credit card fraud detection difficulty. In this method, original transaction data fraud density was used as a trust value and the weighted fraud score was created to reduce the number of misdetections. Fan et al.[13] proposed an application for distributed data mining in the detection of credit card fraud. Brause et al. [14] have also developed an approach that includes advanced data mining techniques and algorithms for the neural network to identify high fraud or cover fraud. Web services and data mining techniques have been proposed by Chiu and Tsai[15] to establish a security system for fraud detection in the banking sector. In this scheme, banks share details in a decentralized environment about the trends of fraud. Web services techniques such as XML, WSDL, etc. are used to create a good channel of data exchange. Phua et al.[16] conducted a survey of existing FDSs based on data mining and published a full report. To detect fraud in credit card transactions, Prodromidis and Stolfo[17] use an agent-based technique with distributed learning. It is based on artificial intelligence and to achieve high accuracy, it combines inductive learning algorithms and meta-learning methods. Phua[18] suggests that meta classifiers should be used as in problems of fraud detection. They regard naïve neural networks of Bayesian, C4.5, and Back Propagation as the base classifiers. A meta classifier is used to determine which classifier should be taken into consideration based on data skew. As a targeted program, they don't specifically use credit card fraud identification, their scheme was generic. Recently, Vatsa et al.[19] suggested a game-theoretical approach to detect credit card fraud. An attacker and an FDS are a match between two players in this strategy in which both seek to maximize their payoff. The problem with most of the methods listed above is that they need marked data to train the classifiers for both legitimate and fraudulent transactions.

3. Hidden Markov Model

A Hidden Markov Model is a finite set of states in which each state is associated with a distribution of probability. Transition probabilities, a set of probabilities govern transitions between these states. A possible outcome or observation that is associated with the image of probability distribution observation is produced in a particular state. It is only the outcome, not the state that is visible to an external observer and therefore states are "hidden" to the outside. Hence the name of the model is the Hidden Markov Model. The Hidden Markov Model is, therefore, an easy and perfect solution to detect fraud through credit card transactions. Another important benefit of the HMM-based approach is an extreme decrease in the number of False Positive transactions that a fraud detection system recognizes as malicious even though they are genuine.

There are 3 canonical problems to solve with HMMs described in:

1. Based on the parameters of the model, calculate the probability of a sequence of output. The Forward and Backward algorithms solve this problem.

2. Based on the parameters of the model, find the most likely sequence of (hidden) states that might have generated a given output sequence. The Viterbi algorithm and Posterior decoding solve this problem.

3. Find the most possible set of probabilities for state change and output given an output sequence. Then solve with the algorithm Baum-Welch.

Our goal is to propose a Hidden Markov Model to reduce high-false positives or high-false alarms and thus improve system performance.

In this prediction process, HMM consider mainly three price value ranges such as

Low (l),
Medium (m) and,
High (h).
First, it will be required to find out the transaction amount belongs to a category either it will be in low, medium, or high ranges.

4. Credit Card Fraud Detection Using HMM

To start the operation of credit card transaction processing in terms of an HMM, we start by deciding the observation values in our model. We quantize the x purchase values into M price ranges i.e. V1; V2; . . . VM, forming the observation symbols at the corresponding bank. For each symbol, the actual price range is based on the spending pattern of individual cardholders. A clustering algorithm determined dynamically these price ranges on each cardholder's transaction values. We use Vk, k ¼ 1; 2;……M, to represent both the corresponding price range and the observation symbol. Here, we take three price ranges, namely, low (l), medium (m), and high(h). Our set of observation symbols is, therefore, V ¼ l; m; h making M ¼ 3. For example, let l=(0, 100], m =(100,500], and h=(500, credit card limit]. If there is a transaction performs by a cardholder is of 190, then the observation symbol is "m". A credit cardholder makes various types of acquisition of various sums over a period. One plausibility is to consider the succession of exchange sums and search for deviations in them. However, the sequence of types of purchase is more stable compared to the sequence of transaction amounts. The reason behind that a cardholder always makes purchases according to his/her needs for procuring different items over a time period. This, in turn, generates a sequence of transaction amounts. Each individual transaction amount associated with the corresponding type of purchase. Hence, we consider the transition in the type of purchase as a state transition in our model. The type of each purchase is linked to the line of business of the corresponding merchant. This data about the dealer's line of business isn't known to the giving bank running the FDS. In this manner, the kind of acquisition of the cardholder is escaped the FDS. The arrangement of every single imaginable sort of procurement and, proportionately, the arrangement of every conceivable line of business of dealers shapes the arrangement of concealed conditions of the HMM. It ought to be noted at this phase the line of business of the dealer is known to the obtaining bank since this data is outfitted at the hour of enrollment of a vendor. Likewise, a few traders might bargain in different sorts of wares (For instance, Wal-Mart, K-Mart, or Target sells a huge number of various things). Such kinds of lines of business are considered as miscellaneous, and we don't endeavor to decide the genuine sorts of things obtained in these exchanges. Any presumption about the accessibility of this data with the giving bank and, henceforth, with the FDS, isn't commonsense and, consequently, would not have been substantial.

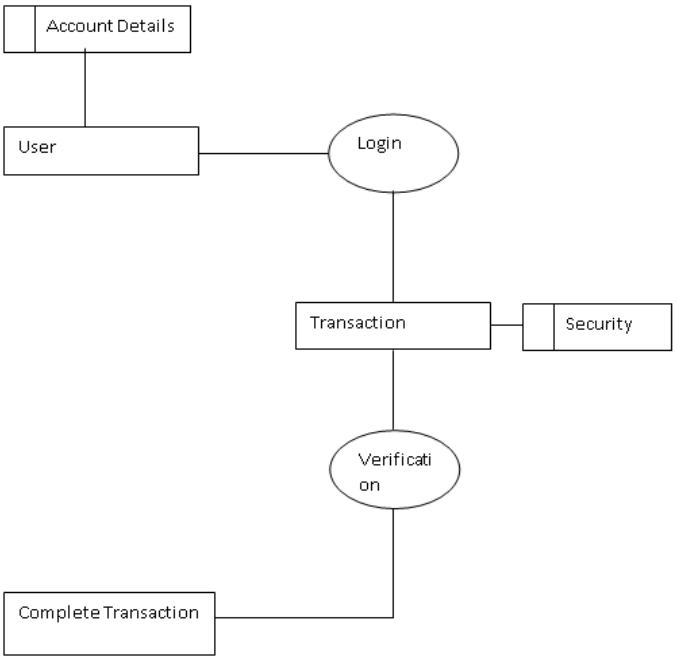

Fig.1: Dataflow diagram of HMM for credit card fraud detection.

5. Proposed System

In the proposed system, we present a Hidden Markov Model (HMM) that don't need fraud signatures and nevertheless is ready to discover frauds by considering a cardholder's payment habit. Transaction sequence by the framework of associate degree HMM. The small print of things purchased in individual transactions unit sometimes not renowned to any Fraud Detection System(FDS) running at the bank that problems credit cards to the cardholders. Hence, we tend to feel that HMM is a perfect alternative for addressing this drawback. Another important advantage of the HMM-based approach may be a forceful reduction within the variety of False Positives transactions known as malicious by associate degree FDS though they're truly real. Every incoming transaction is submitted to the FDS for verification. FDS receives the card details and also the price of the purchase to verify, whether or not the deal is real or not. The categories of products that bought therein dealing don't seem to be renowned for the FDS. It tries to seek out any anomaly within the dealing supported the payment profile of the cardholder, shipping address, and asking address, etc. If the FDS confirms the dealing to be of fraud, it raises associate degree alarm, and also the supply bank declines the transaction.

This new system helps in the detection of the fraud use of the card is faster than the existing system.

In this system no need to check the original user as we maintain a log. The log which is maintained will also be proof for the bank for the transaction made.

We can find the most accurate detection using this technique.

There are two main phases in the proposed system as follows:

5.1 Learning Phase

The learning or training Phase of HMM could be considered into two parts, the first part of this phase converts the transaction amount into observation symbols and form sequences out of them. And the second phase which is often known as the end phase, in this phase we get an HMM for each cardholder. This is the offline step so it will not affect the transaction processing performance, which needs an online response.

5.2 Working Phase

After getting the HMM parameters, we will use symbols for further determination. All symbols which will be opted form costumers training data form an initial sequence of symbols. Let O1, O2.....OR is just an example of such sequence having the length R. This recorded sequence will be formed from the cardholder's transaction up to time t. We will input this sequence to the HMM and measure the probability of acceptance by the HMM.

6. Conclusion

As the credit card transaction process held in many different steps and is represented as the underlying stochastic process of an HMM. It used the all transaction amount as an observation symbol. and the types of items are considered as states of HMM. In this way, we are suggesting a method to find the spending profile of cardholders and then trying to find out the observation symbols which will help us for an initial estimate of the model parameters. It will also be explained how the HMM can detect whether an incoming transaction is genuine or not.

In now a day the online payment mode and fraud in being increased rapidly. So, we must have required specific security technology who can overcome such kind of challenges so that no one can steal money. In this paper, we have observed various approaches already present and tried to propose an updated system that is more useful in these circumstances for detecting Credit Card Frauds.

7. References

[1] V.Bhusari, S.Patil, 2016, "Study of Hidden Markov Model in Credit Card Fraudulent Detection".

[2] Trends in online shopping, a Global Nelson Consumer Report, (2008).

[3] European payment cards fraud report, Payments, Cards and Mobiles LLP & Author, (2010)

[4] International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 7, July 2012)511 Credit Card Fraud Detection Using Hidden Markov Model; Vaibhav Gade, Sonal Chaudhari; All Saint College of Technology, Bhopal (M.P.), India.

[5] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. IEEE, vol. 77, no. 2, pp. 257-286, 1989.

[6] CREDIT CARD FRAUD DETECTION USING HIDDEN MARKOV MODEL by Divya.Iyer, Arti Mohanpurkar, Sneha Janardhan, Dhanashree Rathod, Amruta Sardeshmukh; Department of Computer engi-

neering and Information Technology, MMIT, Pune, India.

[7] Credit Card Fraud Detection Using Hidden Markov Model by Abhinav Srivastava, Amlan Kundu, Shamik Sural, Senior Member, IEEE, and Arun K. Majumdar, Senior Member, IEEE

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

[9] 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).

[10] 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.

[11] Aleskerov, E., Freisleben, B., and Rao, B., 1997. CARD WATCH: 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.

[12] M.J. Kim and T.S. Kim, "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection," Proc.Int'l Conf. Intelligent Data Eng. and Automated Learning, pp. 378-383, 2002.

[13] 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.

[14] R. Brause, T. Langsdorf, and M. Hepp, "Neural Data Mining for Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Tools with Artificial Intelligence, pp. 103-106, 1999.

[15] C. Chiu and C. Tsai, "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection," Proc. IEEE Int'l Conf. e-Technology, e-Commerce and e-Service, pp. 177-181, 2004.

[16] C. Phua, V. Lee, K. Smith, and R. Gayler, "A Comprehensive Survey of Data Mining-Based Fraud Detection Research," https://www.bsys.monash.edu.au/people/cphua/, Mar. 2007.

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

[18] C. Phua, D. Alahakoon, and V. Lee, "Minority Report in Fraud Detection: Classification of Skewed Data," ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 50-59, 2004.

[19] V. Vatsa, S. Sural, and A.K. Majumdar, "A Game-theoretic Approach to Credit Card Fraud Detection," Proc. First Int'l Conf. Information Systems Security, pp. 263-276, 2005

[20] 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.

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

[22] D. Ourston, S. Matzner, W. Stump, and B. Hopkins, "Applications of Hidden Markov Models to Detecting Multi-Stage Network Attacks," Proc. 36th Ann. Hawaii Int'l Conf. System Sciences, vol. 9, pp. 334-344, 2003.

[23] 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.

[24] 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