Second International Conference on Nexgen Technologies
Sengunthar Engineering College, Tiruchengode, Namakkal Dist. Tamilnadu (India)
8th - 9th March 2019 www.conferenceworld.in ISBN : 978-93-87793-75-0

DETECTION OF FRAUD USING TRANSACTION BEHAVIOUR IN CREDIT CARD

Deepika.V[1], Gokila.S[2], Janani.E.R.G[3], Kanagaraju.P[4]

Dept. of Computer Science and Engineering, K.S.Rangasamy College of Technology, Tiruchengode (India)

ABSTRACT

The proliferation of the EMV (Euro-pay MasterCard VISA) chip card design in the credit card business mostly resolved the problem posed by the old Magnetic stripe card technology. However, several works are starting to question the design and implementation of the EMV. This work is suggesting that a detection model must be available to capture the possible anomalous transactions, a fallback in case the technology will fail. One method is to make full use of the historical transaction data including normal transactions and fraud ones to obtain normal/fraud behaviour features based on machine learning techniques, and then utilize these features to check if a transaction is fraud or not. In this project, two kinds of random forests are used to train the behaviour features of normal and abnormal transactions .The proposed system make a comparison of the two random forests which are different in their base classifiers, and analyze their performance on credit fraud detection. Benefits of implementing such detection system will lesser the phone and SMS costs shouldered by the banks; instead of sending SMS transaction notifications to all customers, message will be sent to those customers with detected anomalous transaction.

1. INTRODUCTION

The intention of this study is to fully explore the effectiveness of utilizing the credit card transaction logs to differentiate anomalous from legitimate transactions. The use of credit cards is prevalent in modern day society. But it is obvious that the number of credit card fraud cases is constantly increasing in spite of the chip cards worldwide integration and existing protection systems. This is why the problem of fraud detection is very important now.

In this system two modules (FDS ONLINE and FDS OFFLINE) for fraud detection (transaction classification) are used. The FDS ONLINE module is used for online fraud detection, i.e. fraud detection process during authorization of transactions in a bank processing system.

For the storage of incoming transactions, statistical data for corresponding models, results of classification and generic parameters a FDS Data Warehouse is used. Module FDS ALERT is used for alerting credit card holders in case of fraud recognition by the FDS ONLINE module using SMS or email messages.

The results of the Naïve Bayes method classification using the normal distribution, the kernel density estimation and the discrete distribution are not acceptable for this type of fraud detection. The most legal probabilities for the Naive Bayesian Classifier using these probability estimation methods are too low and are therefore incorrect.

EMV (Euro-pay MasterCard Visa) is a globally accepted standard for chip card based payment transactions, which benefits from the intrinsic security characteristics of chip cards. In existing work described an OPV process that is based on publicly available information. The architecture of the payment network for the OPV and subsequently for the online transaction authorization was explained. This payment network and its associated deployment open up a potential route for an adversary to compromise it to his benefit.

The efficient database management systems have been very important assets for management of a large corpus of data and especially for effective and efficient retrieval of particular information from a large collection whenever needed.

Traditional decision tree classifiers work with data whose values are known and precise. The system extend such classifiers to handle data with uncertain information, which originate from measurement/quantization errors, data staleness, multiple repeated measurements, etc. Although this novel techniques are primarily designed to handle uncertain data, they are also useful for building decision trees using classical algorithms when there are tremendous amounts of data tuples.

2. METHODOLOGY

2.1 CHIP AND PIN IS BROKEN

In this work, Steven J has proposed public policy implications, in light of growing reports of fraud on stolen EMV cards. Frequently, banks deny such fraud victims a refund, asserting that a card cannot be used without the correct PIN, and concluding that the customer must be grossly negligent or lying. Our attack can explain a number of these cases, and exposes the need for further research to bridge the gap between the theoretical and practical security of bank payment systems. It also demonstrates the need for the next version of EMV to be engineered properly. EMV was heavily promoted under the "Chip and PIN" brand during its national rollout in the UK. The technology was advertised as a solution to increasing card fraud: a chip to prevent card counterfeiting, and a PIN to prevent abuse of stolen cards.

2.2 UNDERSTANDING CREDIT CARD FRAUDS

In this work, Tej Paul Bhatla has proposed Credit Card Fraud is one of the biggest threats to business establishments today. However, to combat the fraud effectively, it is important to first understand the mechanisms of executing a fraud. Credit card fraudsters employ a large number of modus operandi to commit fraud. In simple terms, Credit Card Fraud is defined as: When an individual uses another individuals' credit card for personal reasons while the owner of the card and the card issuer are not aware of the fact that the card is being used. Further, the individual using the card has no connection with the cardholder or issuer, and has no intention of either contacting the owner of the card or making repayments for the purchases made. Increasingly, the card not present scenario, such as shopping on the internet poses a greater threat as the merchant (the web site) is no longer protected with advantages of physical verification such as signature check, photo identification, etc .

2.3 CREDIT CARD FRAUD STATISTICS

In this work, Statistic Brain Research Institute has proposed In this study, we initiated the discussion by describing the current situation of the credit card business with respect to fraud issues. Although new technology is available and widely supported by banks and merchants globally to lessen or perhaps eradicate the repercussions of credit card fraud, some researchers are starting to challenge its design and implementation. This work suggests building a model based on the spending behaviour of the card holders and using it to detect anomalous transactions. This work did not elaborate the details of the created model due to a non-disclosure agreement (NDA) between the participating bank and the proponent. However, this work was able to showcase the techniques and processes utilized in building the model. The proponent hopes that in the near future, this work will be used as a reference by some banks or individuals to implement fraud detection system in the financial sector. Benefits of implementing such detection system will lessen the phone and SMS costs shouldered by the banks; instead of sending SMS transaction notifications to all customers, message will be sent to those customers with detected anomalous transaction.

2.4 RESEARCH ON CREDIT CARD FRAUD DETECTION MODEL BASED ON DISTANCE SUM

In this work, Wen-Fang Yu has proposed Using graphs as to extracting and presenting data has a wide range of applications. Such applications may appear in detecting semantic and structural patters and exploiting graphs toward such

applications have steadily been growing. In this work we are going to display one of the most perilous abnormalities in credit cards industry on such concept basis. With advancing technology in field of banking, the rate of use of credit cards has remarkably been escalated. Correspondingly frauds frequency have increased in this area which to surmount such anomalies we model them by means of graphs. Of the prominent advantage of proposed approach is drop of system overload rate during running computations in order to detecting frauds and consequently acceleration of detection speed.

3. PROPOSED SYSTEM

The data pertaining to the fraud cases were collected and encoded in a user defined worksheet while the transaction logs attributes have some similarities to ISO 8583 standard. most point-of-sale (POS) devices and card issuers conform to ISO 8583 – a standard for financial transaction card originated messages. Some attributes of the ISO 8583 standard are present in the transaction logs, although most card systems and credit card companies affix few important details to the log-file. Data annotation will commence by comparing the recorded fraud cases to the transaction logs. When a match is detected, the record in transaction log will have a class value of "true", otherwise "false".

The proposed system says that first, the dataset from the fraud case have to be read from the fraud case file and then it is compared with the complete details of the transaction logs. If the details of the both transactions are matched, it is detected as 'true' and the fraud flag will be turned to 'true'. If it does not matched, then the detected will be turned to 'false' and it goes to the end of the line. Here again check the details if it is 'true' then it goes to the fraud flag that is equal to 'false'. If it detects 'false' then it again goes to the fraud case file to block the fraud person that matches with fraud file.

Data annotation will commence by comparing the recorded fraud cases to the transaction logs . When a match is detected, the record in transaction log will have a class value of "true", otherwise "false".

The fraud detection system says that first, the dataset from the fraud case have to be read from the fraud case file and then it is compared with the complete details of the transaction logs. If the details of the both transactions are matched, it is detected as 'true' and the fraud flag will be turned to 'true'. If it does not matched, then the detected will be turned to 'false' and it goes to the end of the line. Here again check the details if it is 'true' then it goes to the fraud flag that is equal to 'false'. If it detects 'false' then it again goes to the fraud case file to block the fraud person that matches with fraud file.

3.1 DATA ANALYSIS

At the time of this writing, the participating bank had several million of transactions and several thousands of reported fraud cases; however, this work only utilized data from January to June of 2016. Evaluating the conditions of these dataset, it was obviously much skewed – the numbers of legitimate transactions versus fraudulent instances were way too far. According to the research, the preferred distribution must be 50:50 to be able to produce the ideal model. With this considered, and to eliminate the issue of data imbalance, a new dataset was created to contain legitimate transactions based from the card/account included in the recorded fraud case report.

The newly generated dataset has a total count of 9,992 records; 9,733 of which are legitimate and 259 are fraudulent. Although this setup comprises of 3% fraudulent and 97% legitimate transactions, later, this dataset will be manipulated to formulate new datasets with different class distribution – this will confirm the effectiveness of the classifiers under evaluation. For the meantime, the dataset for the model creation and evaluation stage will use a 25% fraud and 75% normal transaction type concoction.

Data in the transaction logs underwent several preprocessing tasks such as datasanitation, normalization, binning, and handling null values. Prior to these activities, few attributes were removed from the dataset such as:

• Account number, card number - this will eliminate the possibility of having a customer centric model. • Fields pertaining to dates - this will reduce the possibility of building a model focusing on date related events. • The control number pertaining

to reported disputes - this will eliminate the possibility of creating a model directly referencing to this control number – since this field represents a potential fraud instance.

4. SYSTEM MODULES

4.1 SECURITY INFORMATION MANAGEMENT

In this module the user has unique card verification value (CCV) and One Time Password (OTP). For login, user is required to enter correct card number and password . In this module the security is provided to the cardholder's data and the transaction that he did. In this module the randomly questions are asked to the cardholder and then the answer given by the cardholder are compared to the database. If the comparison is true then only the cardholder is able to do the transaction. This provides the additional secure layer for transaction.

4.2 TRANSACTION BEHAVIOUR SPENDING PROFILE

This module deals with the overall profile of the user which is related to the transaction. This module generally includes the average amount per transaction, average daily spending, and times of using card and so on. This record is compared with the current transaction which helps for verification.

Apply analytical tools to determine how culture, personality and lifestyle affect the behaviour of consumers and self Analyze how consumer attitudes are formed, and consumer opinions change in order to better understand your own and other's personal consumption experiences Evaluate how different sources of group influences can affect and radically change consumers' consumption Create specific marketing strategies that focus around what motivates consumers, what captures their attention and what retains their loyalty.

4.3 TRANSACTION VERIFICATION AND MONITORING

Verification module verifies whether user entered his data correctly or not. This enables the system to detect whether user is genuine or fraudster .This module is performed after the verification and if the user is not the fraudster then only the transaction is performed according to the cardholder want.

In this module that doing the analysis of different techniques of fraud detection which is user friendly and secure. This system analyzes the feasibility of credit card fraud detection based on outlier mining, applies outlier detection mining based on distance sum into credit card fraud detection and proposes this detection procedures and its empirical process.

4.4 CUSTOMER APPROVAL

The customer places an order on the merchant's website after selecting and adding items in the shopping cart. The Customer details are sent from EBS to the acquiring bank, the acquiring bank then sends the details for authentication. Then customer has to login to his Bank-Account by providing logging details (i.e. username and password) Then purchased amount of the customer will be displayed. The customer is request to enter his credit card number and card verification value (CVV).

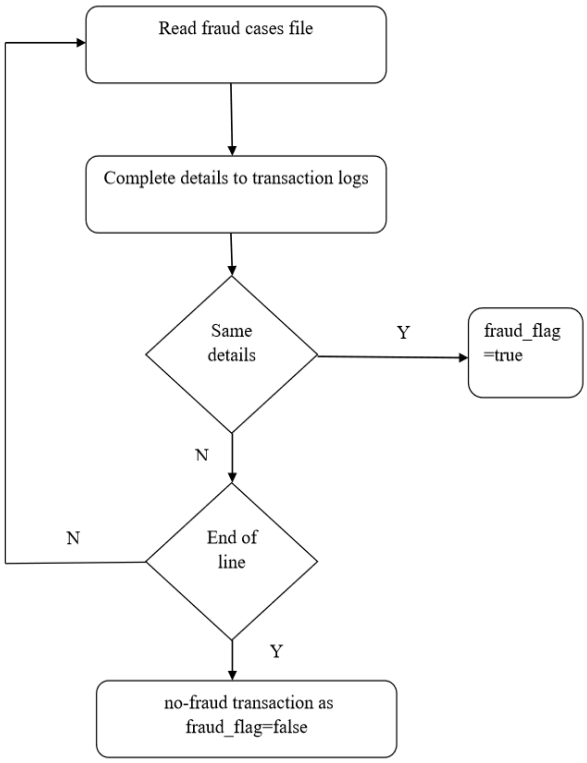

Fig 1: Flow diagram of credit card fraud detection

5. CONCLUSION

In this project, the proposed system initiated the discussion by describing the current situation of the credit card business with respect to fraud issues. Although new technology is available and widely supported by banks and merchants globally to lessen or perhaps eradicate the repercussions of credit card fraud, some researchers are starting to challenge its design and implementation. This work suggests building a model based on the spending behaviour of the card holders and using it to detect anomalous transactions. The proposed work did not elaborate the details of the created model due to a non-disclosure agreement (NDA) between the participating bank and the proponent.

However, the proposed work was able to showcase the techniques and processes utilized in building the model. The proponent hopes that in the near future, the proposed work will be used as a reference by some banks or individuals to implement fraud detection system in the financial sector. Benefits of implementing such detection system will lesser the phone and SMS costs shouldered by the banks; instead of sending SMS transaction notifications to all customers, message will be sent to those customers with detected anomalous transaction.

6. ACKNOWLEDGEMENT

We acknowledge DST-File No.368. DST – FIST (SR/FIST/College – 235/2014 dated 21-11-2014) for financial support and DBT – STAR – College-Scheme-ref.no: BT/HRD/11/09/2018 for providing infrastructure support.

REFERENCE

[1] Beck A. and Teboulle M. (2016), "A fast iterative shrinkage-thresholding algorithm for linear inverse problems", International Journal of Credit Card Fraud Detection, Vol. 2, no. 1, pp. 183–202.

[2] Blitzer J., Dredze M. and Pereira F. (2016), "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification", Jounal of Computer Engineering, Vol. 7, no. 2, pp. 440–447.

[3] Bollen J., Mao H. and Pepe A. (2014), "Modelling public mood and emotion: Twitter sentiment and socio-economic phenomena", Journal of Weblogs Social Media, Vol. 3, no. 4, pp. 17–21.

[4] Boyd S., Parikh N., Chu E., Peleato B. and Eckstein J. (2014), "Distributed optimization and statistical learning via the alternating direction method of multipliers", Found Trends Machine Learning, Vol. 3, no. 1, pp. 1–122.

[5] Cambria E. (2016), "Affective computing and sentiment analysis", IEEE Intelligence System, Vol. 31, no. 2, pp. 102–107.

[6] Cambria E., Schuller B., Xia Y. and White B. (2016), "New avenues in knowledge bases for natural language processing", IEEE Knowledge Based System, Vol. 108, no. 5, pp.1–4.

[7] Chen T., Xu R., He Y., Xia Y., and Wang X. (2013) "Learning user and product distributed representations using a sequence model for sentiment analysis", IEEE Computer Intelligence Magazine, Vol. 11, no. 3, pp. 34–44.

[8] Glorot X., Bordes A., and Bengio Y. (2014) "Domain adaptation for large-scale sentiment classification: A deep learning approach", Journal Machine Learning, Vol. 34, no. 7, pp. 513–520.

[9] Go A., Bhayani R. and Huang L. (2013) "Twitter sentiment classification using distant supervision" .

[10] Hu M .and Liu B. (2014) "Mining and summarizing customer reviews", Journal Knowledge Discovery Data Mining, Vol. 78, no. 6, pp. 168–177.