ISSN:0975-9646

Transaction-Level Behavior Based Credit Card Fraud Detection Mechanism

Bhakti Ratnaparkhi[1], Rahul Patil[2]

[1,2] *ME Computer, PCCOE,*
Pune, India.

Abstract- Now a day technology is increasing very rapidly, which can be used for good as well as for bad purposes. E-commerce is part of all most every system where online transitions are performed through internet through which frauds can be easily done. Credit card system is most vulnerable for frauds. Hence it is very much essential to have fraud detecting system. Till date various approaches have been found by many of the researchers from this area. In this paper we have proposed and implemented new approach by studding various other techniques, their advantages and limitations. It keeps watch on behavior of every user based on which online checking will be done. Comparative results show that performance is improved in our system by avoiding money loss as well as reduced false alarm generation.

***Keywords*- Behavior based, Credit card, fraud detection, Data mining**

I. INTRODUCTION

Fast technology improvement is causing increase in frauds in E-commerce field [1]-[5]. It has been seen that many fraud happen in Credit card systems [3]. Credit card fraud detection is very difficult task. To detect fraudulent transaction it must have some abnormal pattern into it. Transaction which is exactly similar to normal transaction is very hard to be detected.

Credit card fraud has two types of loss. Tangible and intangible, both losses are experienced due to credit card fraud. Loss of money comes under tangible loss which can be said as direct loss but when customer experiences fraud they tell story to many others. Such bad publicity hampers bank's reputation. Bank may lose customer because of such incidences. This is nothing but intangible or indirect loss of bank.

Fraud detection can be broadly divided into two types that are Proactive and Reactive [1], [6]. In Reactive fraud detection mechanism, fraud occurs then records will be scan to detect fraudulent transaction's record. Whereas Proactive mechanism, don't allow fraudulent transactions to get completed successfully. In this mechanism some action (e.g. generation of alarm) will be taken before completion of transaction so that money loose is avoided.

TABLE I
COMPARISON BETWEEN PROACTIVE AND REACTIVE

Proactive	**Reactive**
Action is taken before fraud happens	Action is taken after fraud happens
Fraud is not allowed to happen	Fraud is not allowed to happen then it is detected
Money loss is avoided	Money may get lost
Real time checking	Checking is not real time

In Reactive methods there are two categories as supervised and unsupervised techniques of data mining. All classification techniques come under supervised category where labels of classes are known beforehand. Many algorithms such as DT [7], [9], Bayes Network [8], [11], neural networks [10], support vector machines (SVM), logistic regression, and meta-heuristics such as genetic algorithms are supervised machine learning algorithms [2]. Clustering techniques come under unsupervised category where labels are not known previously and all similar objects are grouped together under one cluster.

Proactive methods are threshold based. Threshold can be maintained globally. But this approach is not found good as it can miss fraudulent transactions having amount less than global threshold, also can generate false alarms for many other transactions. Global threshold idea does not work because it assumes behavior of all customers as same. Better approach will be maintaining threshold based on behavior of customer. Lot of work has been done in this area [1].

TABLE II
COMPARISON BETWEEN GLOBAL THRESHOLD AND BEHAVIOR BASED METHOD

Global Threshold Method	**Behavior Based Threshold Method**
It assumes behavior of all customers as same.	It considers behavior of all customers.
Less accurate.	More accurate.
More false alarms generated.	Fewer false alarms generated.
Easy to implement.	Difficult to implement.

Due to effectiveness of Proactive method we are proposing transaction level fraud detection method here. This method considers behavior of individual customer which will be captured into its signature. In section 2 we will see 2.1model diagram, 2.2initialization of customer behavior, 2.3 Updating behaviors in Signature and 2.4transaction level check. Then we will see discussion in section 3.

II. PROPOSED MODEL

A. Model Diagram

In this paper we are proposing model which performs transaction-level checking as shown in figure 1. Bank can have variety of customers using credit card. To capture their specific behavior we will first classify customers based on their usage of credit card (i.e. amount of money they withdraw from credit card) into three categories like Less, Moderate and High figure 2. Data mining algorithms will be used for this classification. As this is not very

complex classification we will use Decision Tree algorithm to save cost of un-necessary complex calculations.

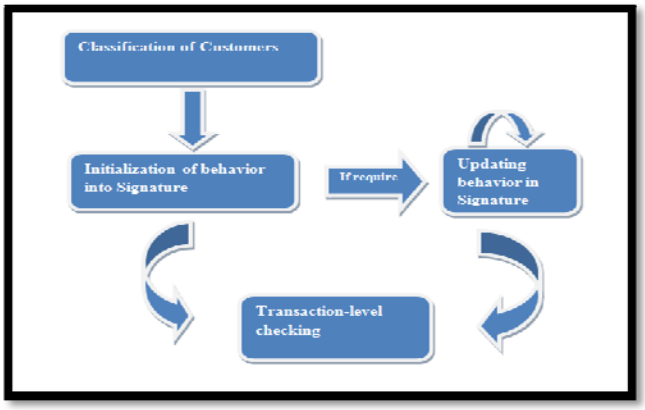

Fig. 1 Model diagram

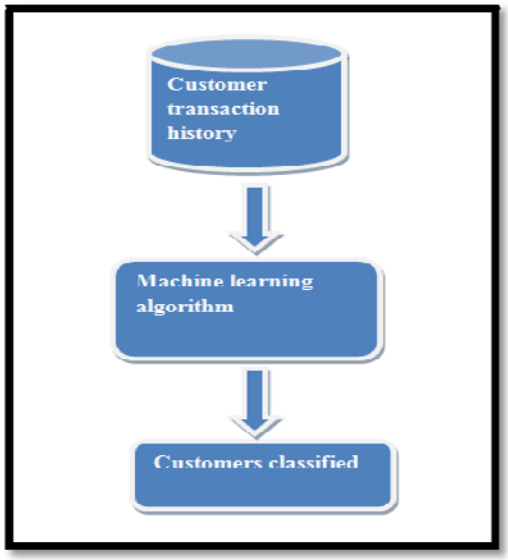

Fig. 2 Customer classification

Following these steps classification rules will be formed which will be in the form:

If X(anticedent) ⟶ Then Y(concequent)

For assessment of rules Coverage and Accuracy can be used:

Coverage= n_covers/|D| (1)

Accuracy= n_correct/n_covers (2)

Where, n_covers is nothing but number of tuples covered by rule R, |D| is training data set and n_correct is number of tuples correctly classified by rule R.

Once customers are classified then signature for customers will be initialized which holds behavior of every individual. Based on the conditions if required this signature will be updated for current behavior. Now every time customer performs transaction using credit card, system will keep watch to make sure that transaction is not fraudulent with the help of captured behavior in individual's signature. If transaction has huge difference than normal behavior of corresponding customer then further action will be taken by holding completion of transaction. This transaction-level checking takes action before fraud happens so is a proactive way of fraud detection. System is divided into 4 parts:

[A] Customer Classification
[B] Initialization of behavior in Signature
[C] Updating behavior in Signature
[D] Transaction-level check

We have seen details of 1st part lets go through others parts of model in detail as well.

B. Initialization of behavior in Signature

To improve the accuracy of fraud detection we need to capture behavior of individual customer. Thus we are using very unique data-structure here to hold behavior of every signal customer which we will be calling it as Signature of individual. Signature will have three fields into it as Date, Threshold (THR), Customer type (CUST_TYP) figure 3.

Date	THR	CUST_TYP

Fig. 3 Signature

Date field will be in the form DD_MM_YY which will indicate how current captured behavior is. In the Signature we are maintaining Threshold value for each customer based on their behavior. This surly improves accuracy as compared to Global threshold mechanism. Last field of Signature indicates Customer type.

As shown in figure 4 current date when Signature is initialized will be placed in first field. Initially THR value will be set to 500Rs as when customer opens account into bank, he is supposed to open it with minimum 500Rs, and so customer may withdraw at most 500Rs at initial level. For setting last field we can take help of classification done in previous part. If customer's type is "Less" then put 1 in this field, 2 for "Moderate" and 3 for "High". It indicates how much money customer used to withdraw using credit card which will help while checking it for fraud. In this way we will initialize customer's behavior into Signature.

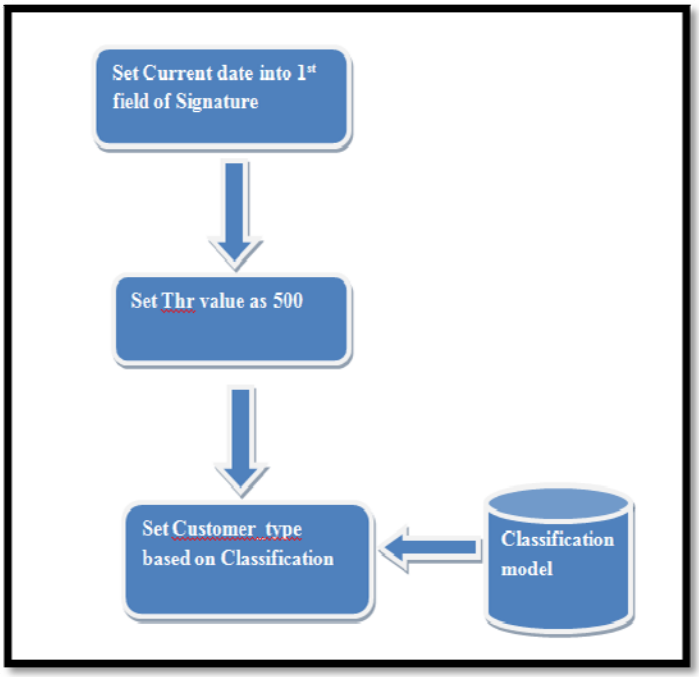

Fig. 4 Initialization Flow

C. *Updating behavior in Signature*

As time passes customer behavior is possible to change. Due many reasons like increment in salary or improvement in standard of living withdraw amount of customer may get increase. To keep watch on this changed behavior of customer Signature of individual should be updated. Figure 5 shows flow of Signature updating activity. After every transaction Signature will be updated with newly calculated values. New THR value will be calculated with the help of transaction history stored into data base as given in equation (3).

$$X = \text{MAX (previous withdrawn amounts)}$$
$$THR = [X + (X*(50/100))] \quad (3)$$

After that current date will placed in first field of signature i.e. Date so that one can keep track of how current is the value in Signature. THR value's validity can last maximum one month period after that forcefully value will be recalculated and validity will be renewed for next month duration. This makes our system as behavior based as we are keeping track of customer's changing behavior by updating Signature.

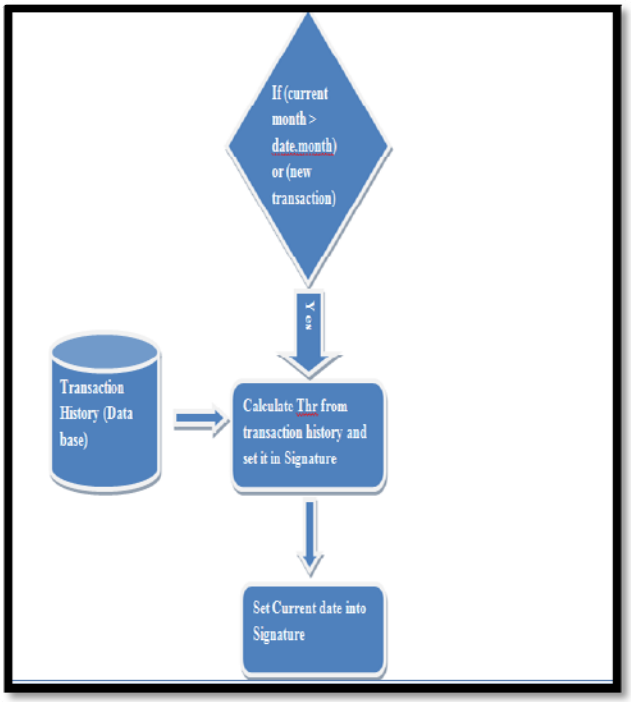

Fig. 5 Flow of Update

D. *Transaction-level check*

This part is responsible for preventing fraudulent transactions to occur. When customer starts transaction by inserting card into system, entered amount will be verified. THR value from his Signature will be extracted to be compared with amount entered. If the amount that customer wants to withdraw is greater than corresponding THR value then transaction can't succeed till second level authentication is performed. Second level authentication is another log in using other password, or alternative question, such as first phone number, first vehicle, favorite restaurant etc. Goal is to verify that correct person have logged in. Steps of transaction-level check are shown bellow.

I. Customer enters card
II. Customer enters amount to be withdrawn
III. Extract customer's Signature
IV. If (Current Month > Month of Signature)
    a. X= MAX (previous withdrawn amounts)
    b. THR= [X + (X*(50/100))]
    c. Date= Current Date
V. If (amount > Signature's THR)
    Second level authentication required
    Else
    Transaction Successful

III. RESULT AND DISCUSSION

We simulated our system and compared outcomes with the system that maintains Global threshold as shown in table 1, 2 and 3. Here we considered 12 customers among which 4 customers are of "Less" type, 4 of "Moderate" and 4 of "High" type. In first case we Global threshold is kept low for more accuracy i.e. 2000. But this will create so many false alarms which decreases performance. Percentage of false alarm generation in this case is: [(8/12)*100] = 66% which is not good. In case 2 Global threshold is kept moderate so percentage of false alarm generation is reduced to [(3/12)*100] = 33% but there are chances of missing frauds for customers of type "Less". In case 3 this threshold is kept high to nullify false alarm but many frauds can be missed in this case. Thus it is very critical task to decide value for Global threshold. Where as in our system we are maintaining THR based on individuals behavior, percentage of false alarm generation is reduced as well as possibility of missing any fraud is also reduced.

Case 1

TABLE III
CASE 1

Name/Id		**Fraud detection in our model**	**Fraud detection in global threshold model** (Global_thr=2000)
(1)	Rupali	Will be detected	Will be detected
(2)	Niket	Will be detected	Will be detected
(3)	Shyam	Will be detected	Will be detected
(4)	Madhura	Will be detected	Will be detected
(5)	Deepali	Will be detected	False alarm
(6)	Lokesh	Will be detected	False alarm
(7)	Aniket	Will be detected	False alarm
(8)	Pritam	Will be detected	False alarm
(9)	Shruti	Will be detected	False alarm
(10)	Suvarna	Will be detected	False alarm
(11)	Bhakti	Will be detected	False alarm
(12)	Supriya	Will be detected	False alarm

Case 2

TABLE IV
CASE 2

Name/Id	Fraud detection in our model	Fraud detection in global threshold model (Global_thr=6000)
(13) Rupali	Will be detected	May/may not be detected
(14) Niket	Will be detected	May/may not be detected
(15) Shyam	Will be detected	May/may not be detected
(16) Madhura	Will be detected	May/may not be detected
(17) Deepali	Will be detected	Will be detected
(18) Lokesh	Will be detected	Will be detected
(19) Aniket	Will be detected	Will be detected
(20) Pritam	Will be detected	Will be detected
(21) Shruti	Will be detected	False alarm
(22) Suvarna	Will be detected	False alarm
(23) Bhakti	Will be detected	False alarm
(24) Supriya	Will be detected	False alarm

Case 3

TABLE V
CASE 3

Name/Id	Fraud detection in our model	Fraud detection in global threshold model (Global_thr=10000)
(1) Rupali	Will be detected	May/may not be detected
(2) Niket	Will be detected	May/may not be detected
(3) Shyam	Will be detected	May/may not be detected
(4) Madhura	Will be detected	May/may not be detected
(5) Deepali	Will be detected	May/may not be detected
(6) Lokesh	Will be detected	May/may not be detected
(7) Aniket	Will be detected	May/may not be detected
(8) Pritam	Will be detected	May/may not be detected
(9) Shruti	Will be detected	Will be detected
(10) Suvarna	Will be detected	Will be detected
(11) Bhakti	Will be detected	Will be detected
(12) Supriya	Will be detected	Will be detected

IV. CONCLUSIONS

In this paper we have provided a system which helps to take precaution against fraudulent transaction in credit card systems. It keeps eye on behavior of every customer and by performing transaction-level checking money loss is avoided. False alarm generation and probability of missing any fraud is also reduced as compared to system which maintains threshold globally. Simulation was performed to show improved performance. Same system can be used for real time environment as a future work.

REFERENCES

[1] Michael Edward Edge, Pedro R. Falcone Sampaio, *A survey of signature based methods for financial fraud detection*, computers & security 28 (2 0 0 9) 3 8 1 – 3 9 4.

[2] Donia Malekian, Mahmoud Reza Hashemi, *An Adaptive Profile based Fraud Detection Framework For Handling Concept Drift*, Information Security and Cryptology (ISCISC), 2013 10th International ISC Conference, IEEE.

[3] *Ekrem Duman , Ayse Buyukkaya , Ilker Elikucuk , A Novel and Successful Credit Card Fraud Detection System Implemented in a Turkish Bank*, 2013 IEEE 13th International Conference on Data Mining Workshops.

[4] Yusuf Sahin, Serol Bulkan, Ekrem Duman, *A cost-sensitive decision tree approach for fraud detection*, Expert Systems with Applications 40 (2013) 5916–5923.

[5] M. Krivko , *A hybrid model for plastic card fraud detection systems*, Expert Systems with Applications 37 (2010) 6070–6076.

[6] Yufeng Kou, Chang-Tien Lu, Sirirat Sinvongwattana , *Survey of Fraud Detection Techniques*, Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control.

[7] Kaiqi Zou , Wenming Sun , Hongzhi Yu , Fengxin Liu , *ID3 decision tree in fraud detection application*, 2012 International Conference on Computer Science and Electronics Engineering.

[8] Yang Guo, Guohua Bai, Yan Hu, *Using Bayes Network for Prediction of Type-2 Diabetes*, 2012, IEEE, 7th International Conference for Internet Technology and Secured Transactions (ICITST).

[9] Ayman Khedr, *Business Intelligence framework to support Chronic Liver Disease Treatment*, International Journal of Computers & Technology Volume 4 No. 2, March-April, 2013, ISSN 2277-3061.

[10] Samuel and Omisore, *Hybrid Intelligent System for the Diagnosis of Typhoid Fever*, J Comput Eng Inf Technol 2013, 2:2, Journal of Computer Engineering & Information Technology.

[11] *Diagnosis of Heart Disease for Diabetic Patients using Naive Bayes Method*, International Journal of Computer Applications (0975 – 8887) Volume 24– No.3, June 2011.