Payment Card Fraud Identification

Rajnish Kumar, Praneet Kumar Gaurav, Swati Shahi, Amol Sitaram Kardel

Abstract— This paper introduces the defensive methods and procedures to identify the payment card fraud. Payment card means credit/debit card which is used for payment purpose over internet. With the rapid advancement in internet, almost all the transaction are being offered by internet as online such as railway ticket booking, mobile recharge, paying the electricity or telephone bill, shopping and etc. this is very good thing because we save our time, we have multi option while shopping but when we transact over internet then chances of fraud also exists. In existing system, we know the fraud happened only when the transaction has been occurred. Sometimes, it become very difficult to identify the fraudulent and hence regarding loses occurs. In this article, we proposed a model namely Advanced Hidden Markov Model which will identify the fraud during transaction. This model is different from Hidden Markov Model. In this Advanced Hidden Markov Model, We used some other set of finite states which is linked through probability distribution states and not visible to user. This model first detect whether it is fraudulent or not and after then it process further so chances of fraud can be minimized using advanced hidden markov model.

Index Terms—PC, AHMM, FIS, FP, PCFI

I. INTRODUCTION

The term Payment Card Fraud Identification deals with the fraud prevention before any transaction from any payment card such as credit card, debit card, charge card and etc. while transacting amount.

In today's life, technology is really the most imperative asset in our lives. It makes our lives so easier and with each passing day we are excessively relying on it for survival. Technology is a central force in today's life. For instance, take an example for communication. earlier we used to keep in touch with people far away through letters which was net by post and it took several days to reach its destination but today through technology a simple email will help us connect with your near and dear ones in just a click.

Now let us take an example for shopping which I done few days earlier. I had to buy a Dell laptop with configuration such as Hard disk drive should be 1 terabyte, processor should be Intel core i7, having graphics card, cd drive, Bluetooth, camera. I went to a nearest dealer shop. I saw all the laptop but configuration was not matching with my requirement after then I searched laptop via internet. Having configuration of my requirement.

Manuscript received on April,2013.

Rajnish Kumar, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, India

Praneet Kumar Gaurav, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, India

Swati Shahi, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, India

Amol Sitaram Kardel, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, India

I found the exact configuration and bought online.in this way, this internet technology provided me to search the laptop according to my choice with respect to going multiple shop but it having also a great disadvantage.

When we are about to pay then sometimes fraud occurs and it becomes very difficult to get regarding loses. If we talk about online transaction then the common thing is payment card (PC).

A payment card is a card which is used to make a payment either by online or offline. Online and offline activities includes paying the electricity/telephone/insurance bill, recharging the mobile/DTH, doing reservation in plane/train/bus. Types of payment card is debit card, credit card, charge card and etc. in below figure, a transaction detail is figured in an online mobile recharge.

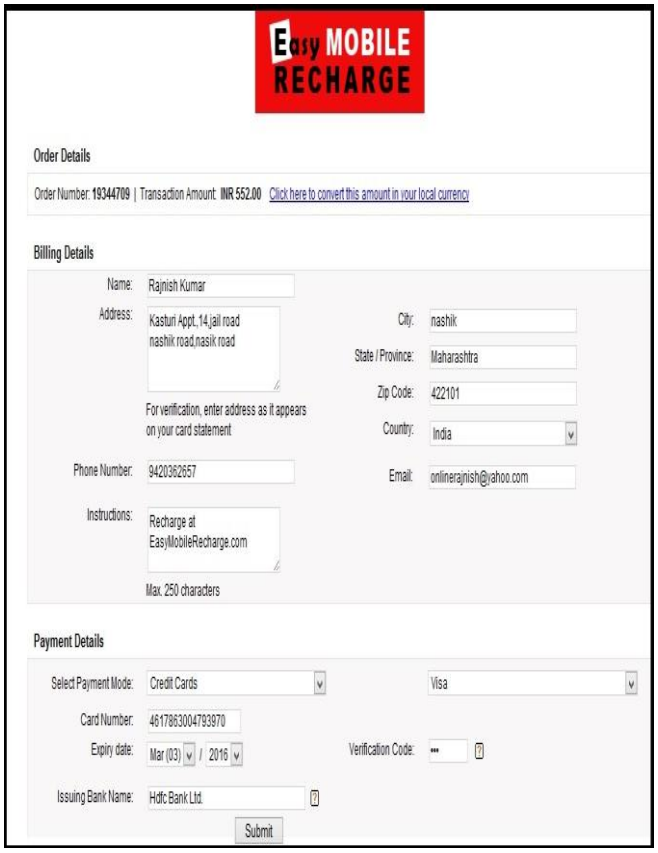

Fig. 1.1 Online Mobile Recharge Using Credit Card

Above figure describes the transaction of amount from an Hdfc Bank Credit Card for online mobile recharge operation. In this figure, all the entry have been made whatever required for mobile recharge. We provide all the credit card details whatever asked. Now it's time to click submit button and processing further.

When we click on submit button then a new window open to verify visa detail. For that we provide same information. Below figure shows the window of authentication.

Fig. 1.2 Verification of Credit Card Details

It verify the credit card details from the issuing bank. Sometimes it fail to verify the details and amount is deducted without performing operation i.e. recharging the mobile number. Now, it's time to complain in bank regarding the loss so that bank will process the required action to come back amount in credit card. It may be time taking but amount can be returned but when we perform online transaction in a unverified website and suppose amount is deducted without performing such operation then it becomes very difficult to get amount so regarding losses occurs.

1.1 Why this Approach Needed?

In existing system, fraud is identified after the transaction been performed which is sometimes difficult to know the fraudulent and minimum chance to get back the amount because when we complain to bank then they first capture the IP address of fraudulent and it leads to a complex cybercrime process.

In this approach, before any transaction performed fraud will be identified. If there won't any fraud in the system then it will perform without any loss and in case of fraud identified, transaction won't be happen.

1.2 Objectives

The main objective of this proposed approach is to automatically detect the internal schema of the system and hide the details for the external user that's why approached name is Advanced Hidden Markov Model. This model uses a finite set of state which is linked through probability distribution and this probability distribution detect the internal schema of that system where the transaction is going to perform.

II. LITERATURE SURVEY

As the payment card having so many facilities though online reservation, recharge, shopping so chances of fraud also be there. A lot of research has been done and a number of techniques has been implemented which mainly focuses on neural network, distributed data mining.

When we talk the payment card fraud identification using neural network then Ghosh and Reilly comes first in mind. Ghosh and Reilly [1] have proposed credit card fraud detection with a neural network. They have built a detection system, which is trained on a large sample of labeled credit card account transactions. These transactions contain example fraud cases due to lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud, and no received issue (NRI) fraud.

Recently, Syeda et al. [2] have used parallel granular neural networks (PGNNs) for improving the speed of data mining and knowledge discovery process in credit card fraud detection. A complete system has been implemented for this purpose.

Stolfo et al. [3] suggest a credit card fraud detection system (FDS) using Meta learning techniques to learn models of fraudulent credit card transactions. Meta learning is a general strategy that provides a means for combining and integrating a number of separately built classifiers or models. A Meta classifier is thus trained on the correlation of the predictions of the base classifiers. The same group has also worked on a cost-based model for fraud and intrusion detection. They use Java agents for Meta learning (JAM), which is a distributed data mining system for credit card fraud detection. A number of important performance metrics like True Positive—False Positive (TP-FP) spread and accuracy have been defined by them.

Aleskerov et al. [4] present CARDWATCH, a database mining system used for credit card fraud detection. The system, based on a neural learning module, provides an interface to a variety of commercial databases.

Kim and Kim [5] have identified skewed distribution of data and mix of legitimate and fraudulent transactions as the two main reasons for the complexity of credit card fraud detection. Based on this observation, they use fraud density of real transaction data as a confidence value and generate the weighted fraud score to reduce the number of misdetections.

Fan et al. [6] suggest the application of distributed data mining in credit card fraud detection.

Brause et al. [7] have developed an approach that involves advanced data mining techniques and neural network algorithms to obtain high fraud coverage.

Chiu and Tsai [8] have proposed Web services and data mining techniques to establish a collaborative scheme for fraud detection in the banking industry. With this scheme, participating banks share knowledge about the fraud patterns in a heterogeneous and distributed environment. Phua et al. [9] have done an extensive survey of existing data-mining-based FDSs and published a comprehensive report.

Prodromidis and Stolfo [10] use an agent-based approach with distributed learning for detecting frauds in credit card transactions. It is based on artificial intelligence and combines inductive learning algorithms and Meta learning methods for achieving higher accuracy.

Vatsa et al. [11] have recently proposed a game-theoretic approach to credit card fraud detection. They model the interaction between an attacker and an FDS as an It is the game between two players, each trying to maximize his payoff. The problem with most of the abovementioned approaches is that they require labeled data for both genuine,

As well as fraudulent transactions, to train the classifiers. Getting real-world fraud data is one of the biggest problems associated with credit card fraud detection. Also, these approaches cannot detect new kinds of frauds for which labeled data is not available.

In contrast, we present a Advanced Hidden Markov Model (AHMM)-based Payment card FIS(Fraud Identification System), which does not require fraud signatures and yet is able to detect frauds by considering a cardholder's spending habit. We model a payment card transaction processing sequence by the stochastic process of an AHMM. The details of items purchased in individual transactions are usually not known to an FIS running at the bank that issues debit/credit cards to the cardholders. This can be represented as the underlying finite Markov chain, which is not observable. The transactions can only be observed through the other stochastic process that produces the sequence of the amount of money spent in each transaction. We will first provide the seed database for structured record after then will extract the raw records from the relevant web page. Once the raw records get extracted then by the use of AHMM parser, matching pattern will be done. Once matching completed, same process will be for unmatched data and merging of new structured data into database will be executed.

III. PROPOSED SYSTEM

In proposed system, we present an Advanced Hidden Markov Model (AHMM) Which does not require fraud signatures and yet is able to detect frauds by considering a cardholder's spending habit. Card transaction processing sequence by the stochastic process of an AHMM. The details of items purchased in Individual transactions are usually not known to any Fraud Identification System (FIS) running at the bank that issues credit cards to the cardholders.

3.1 Advantages

Following are the advantages from this approach:

3.1.1 Reduction in number of FP identified as malicious:

This AHMM-based approach is a drastic reduction in the number of False Positives transactions identified as malicious by an FIS although they are actually genuine. An FIS runs at a credit card issuing bank. Each incoming transaction is submitted to the FIS for verification. FIS receives the card details and the value of purchase to verify, whether the transaction is genuine or not. The types of goods that are bought in that transaction are not known to the FDS. It tries to find any anomaly in the transaction based on the spending profile of the cardholder, shipping address, and billing address, etc. If the FIS confirms the transaction to be of fraud, it raises an alarm, and the issuing bank declines the transaction.

3.1.2 Detect if card used by other in case of card lost

This approach having add an advantage. Suppose my card has been lost and I have booked a complaint regarding my card in the bank then this schema will identify whether others are using my card or not.

When a user login after providing the login details then while performing the transaction, verification is done from both side i.e. from bank and website then a safe transaction is being performed.

IV. ARCHITECTURE

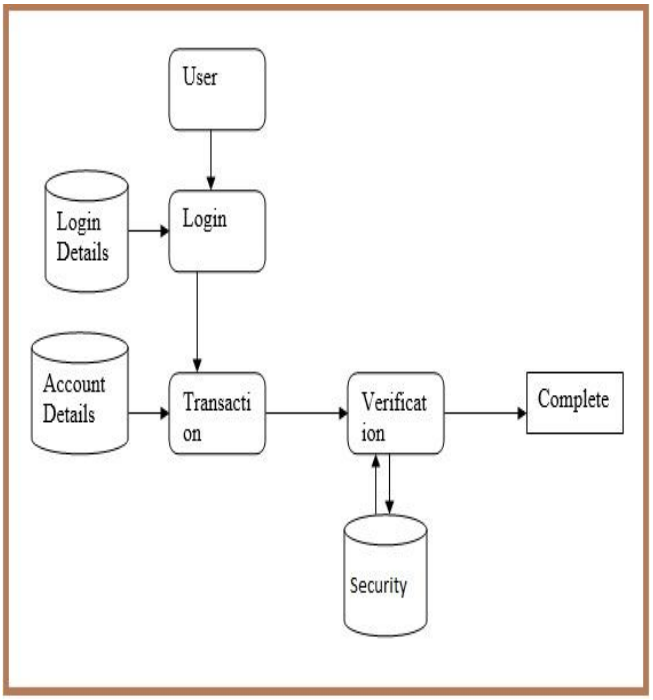

Fig. 4.1 System Architecture

V. APPROACHES USED

This Paper uses the approach of Advanced Hidden Markov Model.

5.1 Advanced Hidden Markov Model

An Advanced hidden Markov model (AHMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (*hidden*) states. An AHMM can be considered the simplest dynamic Bayesian network. The mathematics behind the HMM was developed by L. E. Baum and coworkers. It is closely related to an earlier work on optimal nonlinear filtering problem (stochastic processes) by Ruslan L. Stratonovich, who was the first to describe the forward-backward procedure.

In simpler Markov models (like a Markov chain), the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a *hidden* Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; even if the model parameters are known exactly, the model is still 'hidden'.

Hidden Markov models are especially known for their application in Text Processing, temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics.

A hidden Markov model can be considered a generalization of a mixture model where the hidden variables (or latent variables), which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other.

5.2 Advanced Hidden Markov Model Components

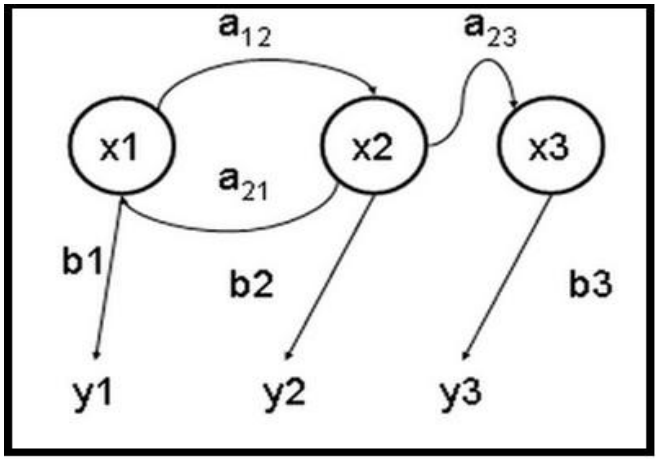

Fig. 5.1 AHMM Components

This Model having Following Components:

- A set of states (x's)
- A set of possible output symbols (y's)
- A state transition matrix (a's)
  - probability of making transition from one state to the next
- Output emission matrix (b's)
  - probability of a emitting/observing a symbol at a particular state
- Initial Probability vector
  - probability of starting at a particular state
  - Not shown, sometimes assumed to be 1

5.3 Advanced Hidden Markov Model Core Problems

There are mainly 3 core problems which has been used in payment card fraud identification:

5.3.1 Evaluation:
Its purpose is to score how well a given model matches a given observation sequence.

5.3.2 Decoding:
Its purpose is to know the hidden states most likely to have generated the observations from a model and a set of observations.

5.3.3 Learning:
Its purpose is to learn AHMM parameters to state transition probabilities and observation probabilities at each state.

5.4 Advanced Hidden Markov Model Concept Summary

We can summarize the concept of this model in following point:

- Build models representing the hidden states of a process or structure using only observations
- Use the models to evaluate probability that a model represents a particular observation sequence
- Use the evaluation information in an application to: recognize speech, parse addresses, and many other applications

5.5 Approach for Payment Card Fraud Identification(PCFI) Using Advanced Hidden Markov Model

Following are the approaches used in Payment Card Fraud Identification:

- Provide Seed Database For Structured Record

Suppose we have to purchase a Camera from www.flipkart.com having price 5000 rupees. First we login in that website either as a guest user or registered user after then we search camera according to our choice. After all things done, when we proceed towards transaction then we have to provide the details of either credit or debit card if we are purchasing the material from card.

Fig. 5.2 Purchasing a Camera From www.flipkart.com

After click Pay button the system will provide seed database for this structured record.

- Extract Raw Records from relevant web pages
  After getting seed database, it will extract the raw records from this flip kart web page.
- Match structured records to raw records
  It will then match the structured records to the raw records.
- Use AHMM Parser
  This model by default having an advanced Hidden Markov Model Parser which is used to scan the structured and raw data.
- Parse unmatched raw recs into structured recs
  The Parser uses match as well as unmatched raw records into structured records.
- Merge new structured records into database

After then it merge the new structured records into the database.

VI. CONCLUSION

In this paper, we have proposed an application of AHMM in payment card Fraud Identification. The different steps in credit card transaction processing are represented as the underlying stochastic process of an AHMM. We have used the ranges of transaction amount as the observation symbols, whereas the types of item have been considered to be states of the AHMM. We have suggested a method for finding the spending profile of cardholders, as well as application of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters. It has also been explained how the AHMM can detect whether an incoming transaction is fraudulent or not. Experimental results show the performance and effectiveness of our system and demonstrate the usefulness of learning the spending profile of the cardholders. Comparative studies reveal that the Accuracy of the system is close to 80 percent over a wide variation in the input data. The system is also scalable for handling large volumes of transactions.

VII. SCOPE FOR FUTURE ENHANCEMENTS

The project has covered almost all the requirements. Further requirements and improvements can easily be done since the coding is mainly structured or modular in nature. Improvements can be appended by changing the existing modules or adding new modules. One important development that can be added to the project in future is file level backup, which is presently done for folder level.

ACKNOWLEDGEMENT

We would like to thank Prof. S. M. Rokade, Head of Computer Engineering Department, Sir Visvesvaraya Institute of Technology, Nasik and our guide Mr. N. Anil, Project Leader, RNC Datacom Solutions Pvt., Ltd., Hyderabad, for his valuable guidance and moral support, without which this paper would not have been possible.
We would also like to thank all other people who have worked earlier on this similar topic as their work helped us a lot.

REFERENCES

[1] Ghosh, S., and Reilly, D.L., 1994. Credit Card Fraud Detection with a Neural-Network, 27th Hawaii International 1 Conference on Information System, vol. 3 (2003), pp. 621- 630
[2] 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)
[3] Stolfo, S. J., Fan, D.W., Lee, W., Prodromidis, A., and Chan, P.K., 2000. Cost-Based Modelling for Fraud and Intrusion Detection: Results from the JAM Project, Proceedings of DARPA Information Survivability Conference and Exposition, Vol. 2 (2000), pp. 130-144.
[4] Aleskerov, E., Freislebeen, 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.
[5] 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.
[6] 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.
[7] 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.
[8] 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.
[9] C. Phau, V. Lee, K. Smith, and R. Gayler, "A Comprehensive Scheme of Data Mining-Based Fraud Detection Research," http://www/bsys.monash.edu.au/people/cphua/, Mar. 2007.
[10] S. Stolfo and A. L. Prodromids, "Agent-Based Distributed Learning Applied to Fraud Detection," Technical Report CUCS-014-99, Columbia Univ., 1999.
[11] V. Vatsa, S. Sral, 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.

Rajnish Kumar is a student in B.E. in the Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, Maharashtra, India.

Praneet Kumar Gaurav is a student in B.E. in the Department of Computer
Engineering, Sir Visvesvaraya Institute Of Technology, Nasik-422101, India

Swati Shahi is a student in B.E. in the epartment of Computer Engineering,
Sir Visvesvaraya Institute of Technology, Nasik-422101, Maharashtra, India

Amol Sitaram Kardel is a student in the Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Nasik-422101, Maharashtra, India