Research Paper *IJBARR*
Impact Factor: 3.072 *E- ISSN -2347-856X*
Peer Reviewed, Listed & Indexed *ISSN -2348-0653*

CREDIT CARD FRAUD DETECTION USING HIDDEN MARKOV MODEL

Dr.P.Anandraj
Assistant Professor, Infant Jesus College of Engineering, Keelavallanadu, Tuticorin.

Abstract
Now a day the usage of credit cards has dramatically increased. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In this paper, we model the sequence of operations in credit card transaction processing using a Hidden Markov Model (HMM) and show how it can be used for the detection of frauds. An HMM is initially trained with the normal behavior of a cardholder. If an incoming credit card transaction is not accepted by the trained HMM with sufficiently high probability, it is considered to be fraudulent. At the same time, we try to ensure that genuine transactions are not rejected. We present detailed experimental results to show the effectiveness of our approach and compare it with other techniques available in the literature.

INTRODUCTION

Credit-card-based purchases can be categorized into two types: 1) physical card and 2) virtual card. In a physical-card based purchase, the cardholder presents his card physically to a merchant for making a payment. To carry out fraudulent transactions in this kind of purchase, an attacker has to steal the credit card. If the cardholder does not realize the loss of card, it can lead to a substantial financial loss to the credit card company. In the second kind of purchase, only some important information about a card (card number, expiration date, secure code) is required to make the payment. Such purchases are normally done on the Internet or over the telephone. To commit fraud in these types of purchases, a fraudster simply needs to know the card details. Most of the time, the genuine cardholder is not aware that someone else has seen or stolen his card information. The only way to detect this kind of fraud is to analyze the spending patterns on every card and to figure out any inconsistency with respect to the “usual” spending patterns. Fraud detection based on the analysis of existing purchase data of cardholder is a promising way to reduce the rate of successful credit card frauds. Since humans tend to exhibit specific behaviorist profiles, every cardholder can be represented by a set of patterns containing information about the typical purchase category, the time since the last purchase, the amount of money spent, etc. Deviation from such patterns is a potential threat to the system. Credit card fraud detection using hidden markov model developed using .net using c#. Modules display as follows.

- New card
- Login
- Security information
- Transaction
- Verification

New card

In this module, the customer gives there information to enroll a new card. The information is all about their contact details. They can create their own login and password for their future use of the card.

Login

In Login Form module presents site visitors with a form with username and password fields. If the user enters a valid username/password combination they will be granted access to additional resources on website. Which additional resources they will have access to can be configured separately.

Security information

In Security information module it will get the information detail and its store’s in database. If the card lost then the Security information module form arise. It has a set of question where the user has to answer the correctly to move to the transaction section. It contain informational privacy and informational self-determination are addressed squarely by the invention affording persons and entities a trusted means to user, secure, search, process, and exchange personal and/or confidential information.

Transaction

The method and apparatus for pre-authorizing transactions includes providing a communications device to a vendor and a credit card owner. The credit card owner initiates a credit card transaction by communicating to a credit card number, and storing therein, a distinguishing piece of information that characterizes a specific transaction to be made by an authorized user of the credit card at a later time. The information is accepted as "network data" in the data base only if a correct personal

identification code (PIC) is used with the communication. The "network data" will serve to later authorize that specific transaction. The credit card owner or other authorized user can then only make that specific transaction with the credit card. Because the transaction is pre-authorized, the vendor does not need to see or transmit a PIC.

Verification

Verification information is provided with respect to a transaction between an initiating party and a verification-seeking party, the verification information being given by a third, verifying party, based on confidential information in the possession of the initiating party. In verification the process will seeks card number and if the card number is correct the relevant process will be executed. If the number is wrong, mail will be sent to the user saying the card no has been block and he can't do the further transaction.

System Specification

Hardware Configuration

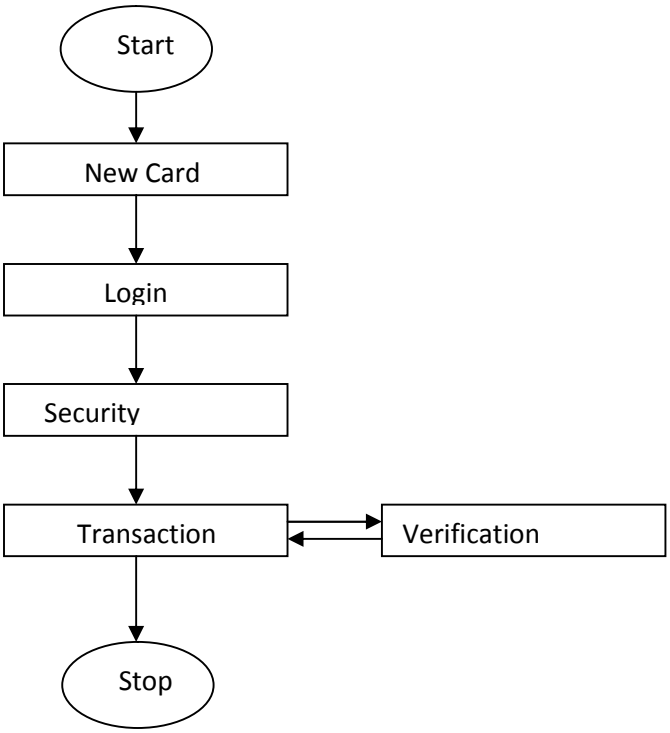
System : Pentium IV 2.4 GHz
Hard disk : 40 GB
Floppy drive : 1.44 MB
Monitor : 15 VGA colour
Mouse : Logitech.
RAM : 256 MB

Software Configuration

Operating system : Windows XP Professional
Front End : Asp .Net 2.0.
Coding Language : Visual C# .Net
Back-End : Sql Server 2000

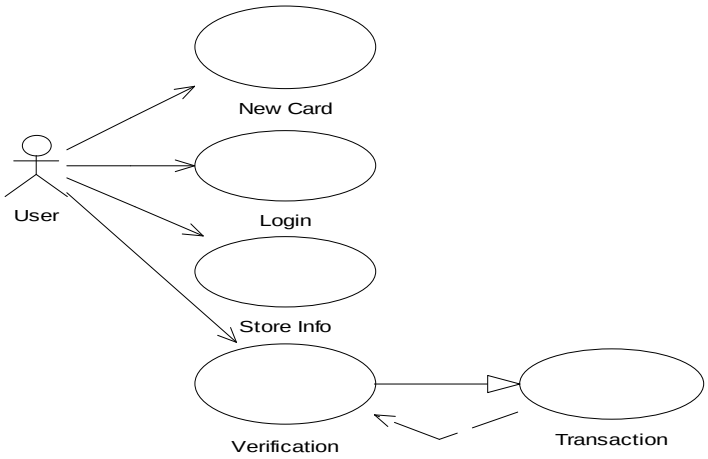
SYSTEM DESIGN

Module diagram

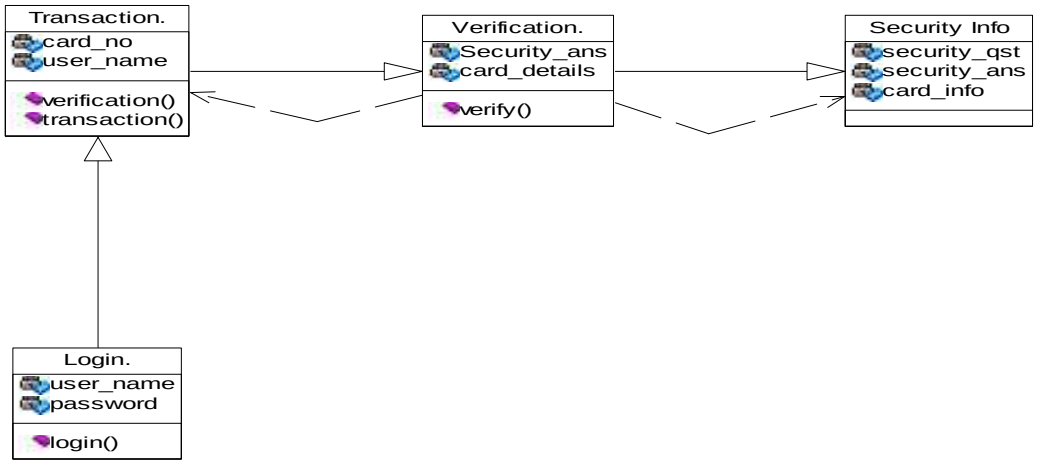


UML Diagrams

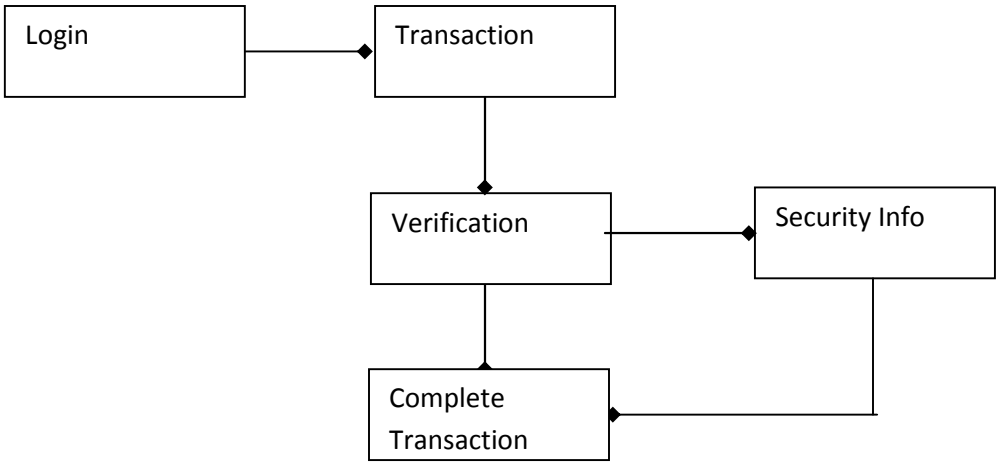
Use case diagram



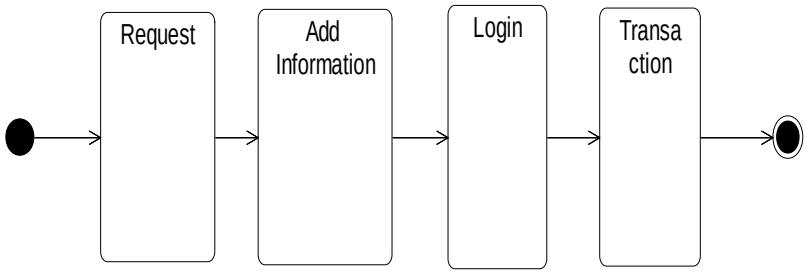
Class diagram



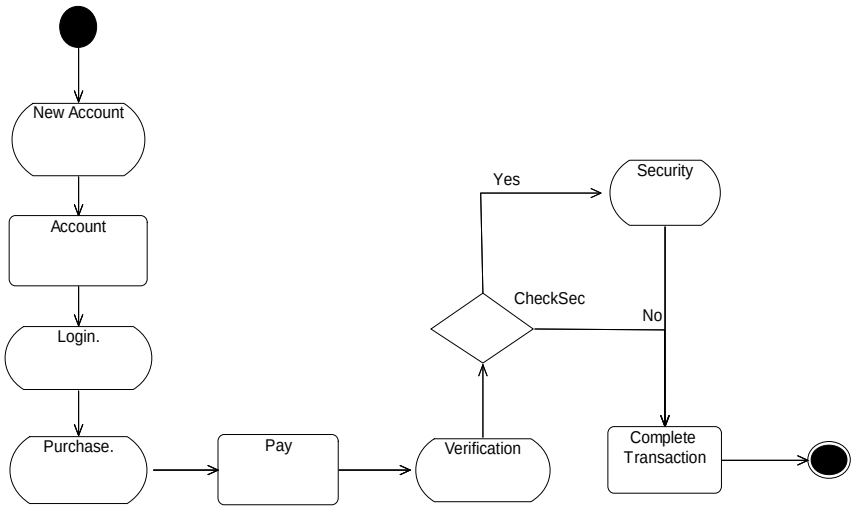
Object diagram



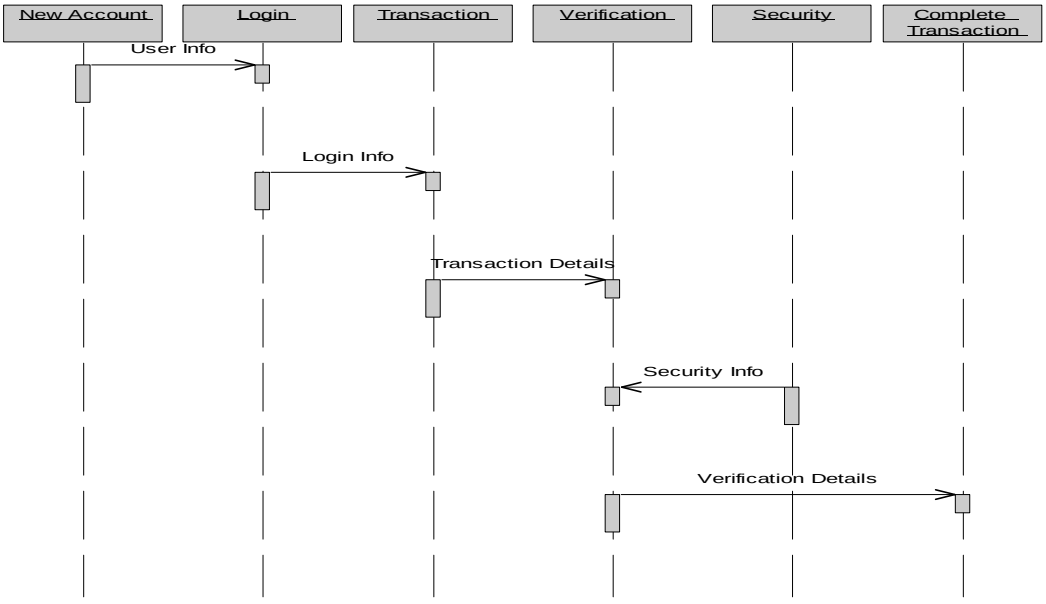
State diagram



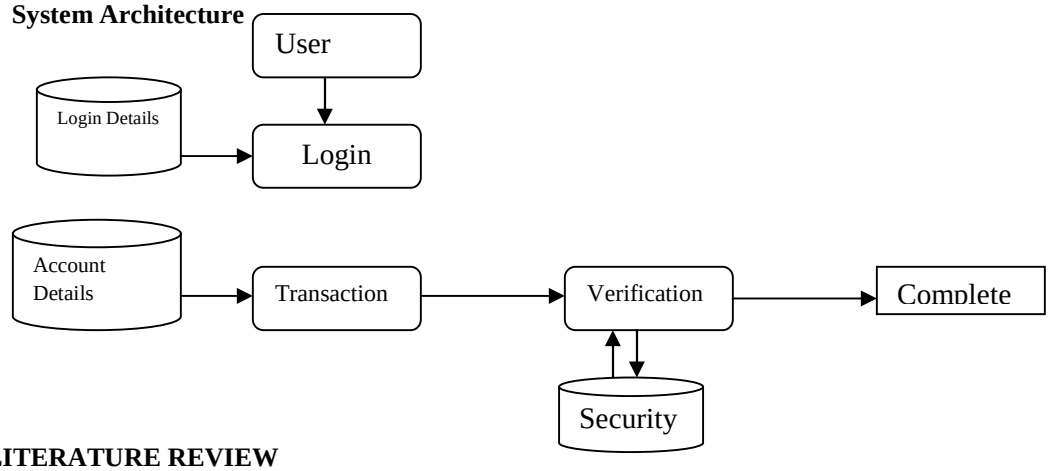
Activity diagram



Sequence diagram



System Architecture



LITERATURE REVIEW

Ghosh and Reilly 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 non received issue (NRI) fraud. Recently, Syeda et al. 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. suggest a credit card fraud detection system (FDS) using metal earning 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. 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. Fan et al. suggest the application of distributed data mining in credit card fraud detection. Brause et al. have developed an approach that involves advanced data mining techniques and neural network algorithms to obtain high fraud coverage. Chiu and Tsai 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. To establish a smooth channel of data exchange, Web services techniques such as XML, SOAP, and WSDL are used. Phua et al. have done an extensive survey of existing data-mining-based FDSs and published a comprehensive report. Prodromidis and Stolfo 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. Phua .suggest the use of meta classifier similar to in fraud detection problems.They consider naiveBayesian, C4.5, and Back Propagation neural networks as the base classifiers. A meta classifier is used to determine which classifier should be considered based on skewness of data. Although they do not directly use credit card fraud detection as the target application, their approach is quite generic. Vatsa et al. have recently proposed a game-theoretic approach to credit card fraud detection. They model the interaction between an attacker and an FDSas a lt is tage 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 Hidden Markov Model (HMM)-based credit card FDS, which does not require fraud signatures and yet is able to detect frauds by considering a cardholder's spending habit.Wemodel a credit card transaction processing sequence by the stochastic process of an HMM. The details of items purchased in individual transactions are usually not known to an FDS running at the bank that issues 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. Hence, we feel that HMM is an ideal choice for addressing this problem. Another important advantage of the HMM-based approach is a drastic reduction in the number of False Positives (FPs)—transactions identified as malicious by an FDS although they are actually genuine. Since the number of genuine transactions is a few orders of magnitude higher than the number of malicious transactions, an FDS should be designed in such a way that the number of FPs is as low as possible.

Techniques and Algorithm Used - HMM Model

To map the credit card transaction processing operation in terms of an HMM, we start by first deciding the observation symbols in our model.Wequantize the purchase values x into M price ranges V1; V2; . . . VM, forming the observation symbols at the issuing bank. The actual price range for each symbol is configurable based on the spending habit of individual cardholders. These price ranges can be determined dynamically by applying a clustering algorithm on the values of each cardholder's transactions, as shown in Section 5.2. We use Vk, k ¼ 1; 2; . . .M, to represent both the observation symbol, as well as the corresponding price range. In this work, we consider only three price ranges, namely, low (l), medium (m), and high(h). Our set of observation symbols is, therefore, V ¼ fl; m; hg making M ¼ 3. For example, let l=(0, $100], m =($100,$500], and h=($500, credit card limit]. If a cardholder performs a transaction of $190, then the corresponding observation symbol is m.

A credit cardholder makes different kinds of purchases of different amounts over a period of time. One possibility is to consider the sequence of transaction amounts and look for deviations in them. However, the sequence of types of purchase is more stable compared to the sequence of transaction amounts. The reason is that, a cardholder makes purchases depending on his need for procuring different types of items over a period of time. This, in turn, generates a sequence of transaction amounts. Each individual transaction amount usually depends on the corresponding type of purchase. Hence, we consider the transition in the type of purchase as state transition in our model. The type of each purchase is linked to the line of business of the corresponding merchant. This information about the merchant's line of business is not known to the issuing bank running the FDS. Thus, the type of purchase of the cardholder is hidden from the FDS. The set of all possible types of purchase and, equivalently, the set of all possible lines of business of merchants forms the set of hidden states of the HMM. It should be noted at this stage that the line of business of the merchant is known to the acquiring bank, since this information is furnished at the time of registration of a merchant. Also, some merchants may be dealing in various types of commodities (For example, Wal-Mart, K-Mart, or Target sells tens of thousands of different items). Such types of line of business are considered as Miscellaneous, and we do not attempt to determine the actual types of items purchasedin these transactions. Any assumption about availability of this information with the issuing bank and, hence, with the FDS, is not practical and, therefore, would not have been valid.

Advantages

- Highly Security from unauthorized use of credit card
- Avoids fraud usage of card through online transactions.
- Detect if card used by others if card lost.

CONCLUSIONS

In this paper, we have proposed an application of HMM in credit card fraud detection. The different steps in credit card transaction processing are represented as the underlying stochastic process of an HMM. 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 HMM. 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 HMM 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.

REFERENCES

1. "Global Consumer Attitude towards On-Line Shopping," http://www2.acnielsen. com/reports/ documents/2005_ cc_online shopping.pdf, Mar. 2007.
2. D.J. Hand, G. Blunt, M.G. Kelly, and N.M. Adams, "Data Mining for Fun and Profit," Statistical Science, vol. 15, no. 2, pp. 111-131, 2000.
3. "Statistics for General and On-Line Card Fraud," http://www. epaynews.com/statistics/fraud.html, Mar. 2007.
4. S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int'l Conf. System Sciences:
5. Information Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.
6. M. Syeda, Y.Q. Zhang, and Y. Pan, "Parallel Granular Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Fuzzy Systems, pp. 572-577, 2002.
7. S.J. Stolfo, D.W. Fan, W. Lee, A.L. Prodromidis, and P.K. Chan, "Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results," Proc. AAAI Workshop AI Methods in Fraud and Risk Management, pp. 83-90, 1997.