CHAPTER 2

LITERATURE REVIEW

2.1 CREDIT CARD FRAUD ANALYSIS

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. Credit card frauds are committed in the following ways:

- An act of criminal deception (mislead with intent) by use of unauthorized account and/or personal information
- Illegal or unauthorized use of account for personal gain
- Misrepresentation of account information to obtain goods and/or services.

Contrary to popular belief, merchants are far more at risk from credit card fraud than the cardholders. While consumers may face trouble trying to get a fraudulent charge reversed, merchants lose the cost of the product sold, pay chargeback fees, and fear from the risk of having their merchant account closed. 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.

In fact, it is almost impossible to perform any of the ‘physical world’ checks necessary to detect who is at the other end of the transaction. This makes the internet extremely attractive to fraud perpetrators. According to a recent survey, the rate at which internet fraud occurs is 12 to 15 times higher than ‘physical world’ fraud. However, recent technical developments are showing some promise to check fraud in the card not present scenario.

While the exact amount of losses due to fraudulent activities on cards is unknown, various research analyst reports concur that the figure for year 2002 probably exceeds $2.5 billion. Further, as the overall e-commerce volumes continue to grow and fraudsters adopt more complex schemes, the projected figure for losses to internet merchants in the US alone is expected to be in the range of $5–15 billion by the year 2005.

This again is dependent on how rapidly fraud prevention technology will be adopted by the industry. The incidence of fraud for credit card transactions taking place over the internet is according to Garner G2, nearly 15 times higher than face-to-face transactions. The increased likelihood of fraud, in conjunction with the full economic liability for fraud losses makes risk management one of the most important challenges for Internet merchants worldwide.

While lost or stolen card is the most common type of fraud, others are listed below:

- Counterfeit
As indicated above, there are many ways in which fraudsters execute a credit card fraud. As technology changes, so do the technology of fraudsters, and thus the way in which they go about carrying out fraudulent activities. Frauds can be broadly classified into three categories, i.e., traditional card related frauds, merchant related frauds and internet frauds.

2.2 CARD RELATED FRAUDS

Application fraud

This type of fraud occurs when a person falsifies an application to acquire a credit card. Application fraud can be committed in three ways:

- Assumed identity, where an individual illegally obtains personal information of another individual and opens accounts in his or her name, using partially legitimate information.
- Financial fraud, where an individual provides false information about his or her financial status to acquire credit.

- Not-received items (NRIs) also called postal intercepts occur when a credit card is stolen from the postal service before it reaches its owner’s destination.

**Lost/ Stolen cards**

A card is lost/stolen when a legitimate account holder receives a card and loses it or someone steals the card for criminal purposes. This type of fraud is in essence the easiest way for a fraudster to get hold of other individual's credit cards without investment in technology. It is also perhaps the hardest form of traditional credit card fraud to tackle.

**Account takeover**

This type of fraud occurs when a fraudster illegally obtains a valid customers’ personal information. The fraudster takes control of (takeover) a legitimate account by either providing the customer’s account number or the card number. The fraudster then contacts the card issuer, masquerading as the genuine cardholder, to ask that mail be redirected to a new address. The fraudster reports card lost and asks for a replacement to be sent.

**Fake and counterfeit cards**

The creation of counterfeit cards, together with lost / stolen cards poses highest threat in credit card frauds. Fraudsters are constantly finding new and more innovative ways to create counterfeit cards. Some of the techniques used for creating false and counterfeit cards are listed below:
1. Erasing the magnetic strip: A fraudster can tamper an existing card that has been acquired illegally by erasing the metallic strip with a powerful electro-magnet. The fraudster then tampers with the details on the card so that they match the details of a valid card, which they may have attained, e.g., from a stolen till roll. When the fraudster begins to use the card, the cashier will swipe the card through the terminal several times, before realizing that the metallic strip does not work. The cashier will then proceed to manually input the card details into the terminal. This form of fraud has high risk because the cashier will be looking at the card closely to read the numbers. Doctored cards are, as with many of the traditional methods of credit card fraud, becoming an outdated method of illicit accumulation of either funds or goods.

2. Creating a fake card: A fraudster can create a fake card from scratch using sophisticated machines. This is the most common type of fraud though fake cards require a lot of effort and skill to produce. Modern cards have many security features all designed to make it difficult for fraudsters to make good quality forgeries. Holograms have been introduced in almost all credit cards and are very difficult to forge effectively. Embossing holograms onto the card itself is another problem for card forgers.

3. Altering card details: A fraudster can alter cards by either re-embossing them — by applying heat and pressure to the information originally embossed on the card by a legitimate card manufacturer or by re-encoding them using computer software that encodes the magnetic stripe data on the card.

4. Skimming: Most cases of counterfeit fraud involve skimming, a process where genuine data on a card’s magnetic stripe is electronically copied onto another. Skimming is fast emerging as the most popular form of credit card fraud.
Employees/cashiers of business establishments have been found to carry pocket skimming devices, a battery-operated electronic magnetic stripe reader, with which they swipe customer’s cards to get hold of customer’s card details. The fraudster does this whilst the customer is waiting for the transaction to be validated through the card terminal. Skimming takes place unknown to the cardholder and is thus very difficult, if not impossible to trace. In other cases, the details obtained by skimming are used to carry out fraudulent card-not-present transactions by fraudsters. Often, the cardholder is unaware of the fraud until a statement arrives showing purchases they did not make.

5. **White plastic**: A white plastic is a card-size piece of plastic of any color that a fraudster creates and encodes with legitimate magnetic stripe data for illegal transactions. This card looks like a hotel room key but contains legitimate magnetic stripe data that fraudsters can use at POS terminals that do not require card validation or verification (for example, petrol pumps and ATMs).

**2.3 MERCHANT RELATED FRAUDS**

Merchant related frauds are initiated either by owners of the merchant establishment or their employees. The types of frauds initiated by merchants are described below:

**Merchant collusion**

This type of fraud occurs when merchant owners and/or their employees conspire to commit fraud using their customers’ (cardholder) accounts and/or personal information. Merchant owners and/or their employees pass on the information about cardholders to fraudsters.

**Triangulation**
The fraudster in this type of fraud operates from a web site. Goods are offered at heavily discounted rates and are also shipped before payment. The fraudulent site appears to be a legitimate auction or a traditional sales site. The customer while placing orders online provides information such as name, address and valid credit card details to the site. Once fraudsters receive these details, they order goods from a legitimate site using stolen credit card details. The fraudster then goes on to purchase other goods using the credit card numbers of the customer. This process is designed to cause a great deal of initial confusion, and the fraudulent internet company in this manner can operate long enough to accumulate vast amount of goods purchased with stolen credit card numbers.

2.4 INTERNET RELATED FRAUDS

The Internet has provided an ideal ground for fraudsters to commit credit card fraud in an easy manner. Fraudsters have recently begun to operate on a truly transnational level. With the expansion of trans-border or 'global' social, economic and political spaces, the internet has become a New World market, capturing consumers from most countries around the world. The most commonly used techniques in internet fraud are described below:

1. Site cloning: Site cloning is where fraudsters clone an entire site or just the pages from which you place your order. Customers have no reason to believe they are not dealing with the company that they wished to purchase goods or services from because the pages that they are viewing are identical to those of the real site. The cloned or spoofed site will receive these details and send the customer a receipt of the transaction via email just as the real company would. The consumer suspects nothing, whilst the fraudsters have all the details they need to commit credit card fraud.
2. **False merchant sites**: These sites often offer the customer an extremely cheap service. The site requests a customer’s complete credit card details such as name and address in return for access to the content of the site. Most of these sites claim to be free, but require a valid credit card number to verify an individual’s age. These sites are set up to accumulate as many credit card numbers as possible. The sites themselves never charge individuals for the services they provide. The sites are usually part of a larger criminal network that either uses the details it collects to raise revenues or sells valid credit card details to small fraudsters.

3. **Credit card generators**: Credit card number generators are computer programs that generate valid credit card numbers and expiry dates. These generators work by generating lists of credit card account numbers from a single account number. The software works by using the mathematical Luhn algorithm that card issuers use to generate other valid card number combinations. The generators allow users to illegally generate as many numbers as the user desires, in the form of any of the credit card formats, whether it be American Express, Visa or MasterCard.

### 2.5 IMPACT OF CREDIT CARD FRAUDS

Unfortunately, occurrences of credit card frauds have only shown an upward trend so far. The fraudulent activity on a card affects everybody, i.e., the cardholder, the merchant, the acquirer as well as the issuer. This section analyses the impact that credit card frauds have on all the players involved in transacting business through credit cards.

**Impact of Fraud on Cardholders**
It's interesting to note that cardholders are the least impacted party due to fraud in credit card transactions as consumer liability is limited for credit card transactions by the legislation prevailing in most countries. This is true for both card-present as well as card-not-present scenarios. Many banks even have their own standards that limit the consumer's liability to a greater extent. They also have a cardholder protection policy in place that covers for most losses of the cardholder. The cardholder has to just report suspicious charges to the issuing bank, which in turn investigates the issue with the acquirer and merchant, and processes chargeback for the disputed amount.

Impact of Fraud on Merchants

Merchants are the most affected party in a credit card fraud, particularly more in the card-not-present transactions, as they have to accept full liability for losses due to fraud. Whenever a legitimate cardholder disputes a credit card charge, the card-issuing bank will send a chargeback to the merchant (through the acquirer), reversing the credit for the transaction.

In case, the merchant does not have any physical evidence (e.g. delivery signature) available to challenge the cardholder's dispute, it is almost impossible to reverse the chargeback. Therefore, the merchant will have to completely absorb the cost of the fraudulent transaction. In fact, this cost consists of several components, which could add up to a significant amount. The cost of a fraudulent transaction consists of:

1. **Cost of goods sold**: Since it is unlikely that the merchandise will be recovered in a case of fraud, the merchant will have to write off the value of goods involved in a
fraudulent transaction. The impact of this loss will be highest for low-margin merchants.

2. **Shipping cost**: More relevant in a *card-not-present* scenario. Since the shipping cost is usually bundled in the value of the order, the merchant will also need to absorb the cost of shipping for goods sold in a fraudulent transaction. Furthermore, fraudsters typically request high-priority shipping for their orders to enable rapid completion of the fraud, resulting in high shipping costs.

3. **Card association fees**: Visa and MasterCard have put in place fairly strict programs that penalize merchants generating excessive charge-backs. Typically, if a merchant exceeds established chargeback rates for any three-month period (e.g. 1% of all transactions or 2.5% of the total dollar volume), the merchant could be penalized with a fee for every chargeback. In extreme cases, the merchant’s contract to accept cards could be terminated.

4. **Merchant bank fees**: In addition to the penalties charged by card associations, the merchant has to pay an additional processing fee to the acquiring bank for every chargeback.

5. **Administrative cost**: Every transaction that generates a chargeback requires significant administrative costs for the merchant. On average, each chargeback requires one to two hours to process. This is because processing a chargeback requires the merchant to receive and research the claim, contact the consumer, and respond to the acquiring bank or issuer with adequate documentation.
6. **Loss of Reputation**: Maintaining reputation and goodwill is very important for merchants as excessive charge-backs and fraud monitoring could both drive cardholders away from transacting business with a merchant.

2.6 **IMPACT OF FRAUD ON BANKS (ISSUER/ACQUIRER)**

   Based on the scheme rules defined by both MasterCard and Visa, it is sometimes possible that the Issuer/Acquirer bears the costs of fraud. Even in cases when the Issuer/Acquirer is not bearing the direct cost of the fraud, there are some indirect costs that will finally be borne by them. Like in the case of charge-backs issued to the merchant, there are administrative and manpower costs that the bank has to incur. The issuers and acquirers also have to make huge investments in preventing frauds by deploying sophisticated IT systems for detection of fraudulent transactions.

2.7 **FRAUD PREVENTION AND MANAGEMENT**

   With all the negative impacts of fraudulent credit card activities – financial and product losses, fines, loss of reputation, etc, and technological advancements in perpetrating fraud – it's easy for merchants to feel victimized and helpless. However, technological advancements in preventing fraud have started showing some promise to combat fraud. Merchants and Acquirers & Issuers are creating innovative solutions to bring down on fraudulent transactions and lower merchant chargeback rates. One of the main challenges with fraud prevention is the long time lag between the time a fraudulent transaction occurs and the time when it gets detected, i.e., the cardholder initiates a chargeback. Analysis shows that the average lag between the transaction date and the chargeback notification could be as high as 72 days. This means that, if no fraud prevention is in place, one or more
fraudsters could easily generate significant damage to a business before the affected stakeholders even realize the problem.

2.8 FRAUD PREVENTION TECHNOLOGIES

While fraudsters are using sophisticated methods to gain access to credit card information and perpetrate fraud, new technologies are available to help merchants to detect and prevent fraudulent transactions. Fraud detection technologies enable merchants and banks to perform highly automated and sophisticated screenings of incoming transactions and flagging suspicious transactions. The various fraud prevention techniques are discussed below:

Manual review

This method consists of reviewing every transaction manually for signs of fraudulent activity and involves an exceedingly high level of human intervention. This can prove to be very expensive, as well as time consuming. Moreover, manual review is unable to detect some of the more prevalent patterns of fraud, such as use of a single credit card multiple times on multiple locations (physical or web sites) in a short span.

Card verification methods

The Card Verification Method3 (CVM) consists of a 3- or 4-digit numeric code printed on the card but is not embossed on the card and is not available in the magnetic stripe. The merchant can request the cardholder to provide this numeric code in case of card-not-present transaction and submit it with authorization. The purpose of CVM is to ensure that the person submitting the transaction is in possession of the actual card, since the code cannot be copied from receipts or
skimmed from magnetic stripe. Although CVM provides some protection for the merchant, it doesn’t protect them from transactions placed on physically stolen cards. Furthermore, fraudsters who have temporary possession of a card could, in principle, read and copy the CVM code.

**Negative and Positive lists**

A negative list is a database used to identify high-risk transactions based on specific data fields. An example of a negative list would be a file containing all the card numbers that have produced charge-backs in the past, used to avoid further fraud from repeat offenders. Similarly a merchant can build negative lists based on billing names, street addresses, emails and internet protocols (IPs) that have resulted in fraud or attempted fraud, effectively blocking any further attempts. A merchant/acquirer could create and maintain a list of high-risk countries and decide to review or restrict orders originating from those countries.

Another popular example of negative list is the SAFE file distributed by MasterCard to merchants and member banks. This list contains card numbers, which could be potentially used by fraudsters, e.g., cards that have been reported as lost or stolen in the immediate recent past. Positive files are typically used to recognize trusted customers, perhaps by their card number or email address, and therefore bypass certain checks. Positive files represent an important tool to prevent unnecessary delays in processing valid orders.

**Payer authentication**

Payer authentication is an emerging technology that promises to bring in a new level of security to business-to-consumer internet commerce. The first
implementation of this type of service is the Verified by Visa (VbV) or Visa Payer Authentication Service (VPAS) program, launched worldwide by Visa in 2002. The program is based on a Personal Identification Number (PIN) associated with the card, similar to those used with ATM cards, and a secure direct authentication channel between the consumer and the issuing bank. The PIN is issued by the bank when the cardholder enrolls the card with the program and will be used exclusively to authorize online transactions. When registered cardholders check out at a participating merchant’s site, they will be prompted by their issuing bank to provide their password. Once the password is verified, the merchant may complete the transaction and send the verification information on to their acquirer.

**Lockout mechanisms**

Automatic card number generators represent one of the new technological tools frequently utilized by fraudsters. These programs, easily downloadable from the Web, are able to generate thousands of ‘valid’ credit card numbers. The traits of frauds initiated by a card number generator are the following:

- Multiple transactions with similar card numbers (e.g. same Bank Identification Number (BIN))

- A large number of declines

Acquiring banks/merchant sites can put in place prevention mechanisms specifically designed to detect number generator attacks.

**Fraudulent merchants**

Both MasterCard and Visa publish a list of merchants who have been known for being involved in fraudulent transactions in the past. These lists (NMAS -
from Visa and MATCH - from MasterCard) could provide useful information to acquirers right at the time of merchant recruitment preventing potential fraudulent transactions.

2.9 RECENT DEVELOPMENTS IN FRAUD MANAGEMENT

The technology for detecting credit card frauds is advancing at a rapid pace – rules based systems, neural networks, chip cards and biometrics are some of the popular techniques employed by Issuing and Acquiring banks these days. Apart from technological advances, another trend which has emerged during the recent years is that fraud prevention is moving from back-office transaction processing systems to front-office authorization systems to prevent committing of potentially fraudulent transactions. However, this is a challenging trade-off between the response time for processing an authorization request and extent of screening that should be carried out.

Simple rule systems

Simple rule systems involve the creation of ‘if...then’ criteria to filter incoming authorizations/transactions. Rule-based systems rely on a set of expert rules designed to identify specific types of high-risk transactions. Rules are created using the knowledge of what characterizes fraudulent transactions. For instance, a rule could look like – If transaction amount is > $5000 and card acceptance location = Casino and Country = ‘a high-risk country’. Fraud rules enable to automate the screening processes leveraging the knowledge gained over time regarding the characteristics of both fraudulent and legitimate transactions. Typically, the effectiveness of a rule-based system will increase over time, as more rules are added.
to the system. It should be clear, however, that ultimately the effectiveness of the system depends on the knowledge and expertise of the person designing the rules.

The disadvantage of this solution is that it can increase the probability of throwing many valid transactions as exceptions; however, there are ways by which this limitation can be overcome to some extent by prioritizing the rules and fixing limits on number of filtered transactions.

**Risk scoring technologies**

Risk scoring tools are based on statistical models designed to recognize fraudulent transactions, based on a number of indicators derived from the transaction characteristics. Typically, these tools generate a numeric score indicating the likelihood of a transaction being fraudulent: the higher the score, the more suspicious the order. Risk scoring systems provide one of the most effective fraud prevention tools available. The primary advantage of risk scoring is the comprehensive evaluation of a transaction being captured by a single number.

While individual fraud rules typically evaluate a few simultaneous conditions, a risk-scoring system arrives at the final score by weighting several dozens of fraud indicators, derived from the current transaction attributes as well as cardholder historical activities. E.g., transaction amounts more than three times the average transaction amount for the cardholder in the last one year. The second advantage of risk scoring is that, while a fraud rule would either flag or not flag a transaction, the actual score indicates the degree of suspicion on each transaction.
Thus, transactions can be prioritized based on the risk score and given a limited capacity for manual review, only those with the highest score would be reviewed.

**Neural network technologies**

Neural networks are an extension of risk scoring techniques. They are based on the ‘statistical knowledge’ contained in extensive databases of historical transactions, and fraudulent ones in particular. These neural network models are basically ‘trained’ by using examples of both legitimate and fraudulent transactions and are able to correlate and weigh various fraud indicators (e.g., unusual transaction amount, card history, etc) to the occurrence of fraud. A neural network is a computerized system that sorts data logically by performing the following tasks:

- Identifies cardholder’s buying and fraudulent activity patterns.
- Processes data by trial and elimination (excluding data that is not relevant to the pattern).
- Finds relationships in the patterns and current transaction data.

The principles of neural networking are motivated by the functions of the brain – especially pattern recognition and associative memory. The neural network recognizes similar patterns, predicting future values or events based upon the associative memory of the patterns it has learned. The advantages neural networks offer over other techniques are that these models are able to learn from the past and thus, improve results as time passes. They can also extract rules and predict future activity based on the current situation. By employing neural networks effectively, banks can detect fraudulent use of a card, faster and more efficiently.
**Biometrics**

Biometrics is the name given to a fraud prevention technique that records a unique characteristic of the cardholder like, a fingerprint or how he/she sign his/her name, so that it can be read by a computer. The computer can then compare the stored characteristic with that of the person presenting the card to make sure that the right person has the right card. Biometrics, which provides a means to identify an individual through the verification of unique physical or behavioral characteristics, seems to supersede PIN as a basis for the next generation of personal identity verification systems. There are many types of biometrics systems under development such as fingerprint verification, hand based verification, retinal and iris scanning and dynamic signature verification.

**Address verification system**

This technique is applicable in card-not-present scenarios. Address Verification System (AVS) matches the first few digits of the street address and the ZIP code information given for delivering/billing the purchase to the corresponding information on record with the card issuers. A code representing the level of match between these addresses is returned to the merchant. AVS is not much useful in case of international transactions.

**Smart cards**

To define in the simplest terms, a smart card is a credit card with some intelligence in the form of an embedded CPU. This card-computer can be programmed to perform tasks and store information, but the intelligence is limited – meaning that the smart card's power falls far short of a desktop computer.
Smart credit cards operate in the same way as their magnetic counterparts, the only difference being that an electronic chip is embedded in the card. These smart chips add extra security to the card. Smart credit cards contain 32-kilobyte microprocessors, which is capable of generating 72 quadrillion or more possible encryption keys and thus making it practically impossible to fraudulently decode information in the chip.

The smart chip has made credit cards a lot more secure; however, the technology is still being run alongside the magnetic strip technology due to a slow uptake of smart card reading terminals in the world market. Smart cards have evolved significantly over the past decade and offer several advantages compared to a general-purpose magnetic stripe card. The advantages are listed below:

- Stores many times more information than a magnetic stripe card.
- Reliable and harder to tamper with than a magnetic stripe card.
- Performs multiple functions in a wide range of industries.
- Compatible with portable electronic devices such as phones and personal digital assistants (PDAs), and with PCs.
- Stores highly sensitive data such as signing or encryption keys in a highly secure manner
- Performs certain sensitive operations using signing or encryption keys in a secure fashion.

A consortium of Europay MasterCard and Visa (EMV) recently issued a set of specifications for embedding chips in credit cards and processing transactions.
from such cards. MasterCard and Visa have also issued deadlines for compliance with these specifications indicating that banks will have to bear a large portion of fraud losses if they do not comply with EMV specifications. However, the market response has been slow so far due to large investments needed in implementing the EMV compliant programs. As card business transactions increase, so too do frauds. Clearly, global networking presents as many new opportunities for criminals as it does for businesses. While offering numerous advantages and opening up new channels for transaction business, the internet has also brought in increased probability of fraud in credit card transactions. The good news is that technology for preventing credit card frauds is also improving many folds with passage of time.

Reducing cost of computing is helping in introducing complex systems, which can analyze a fraudulent transaction in a matter of fraction of a second. It is equally important to identify the right segment of transactions, which should be subject to review, as every transaction does not have the same amount of risk associated with it. Finding the optimally balanced ‘total cost of fraud’ and other measures outlined in this article can assist acquiring and issuing banks in combating frauds more efficiently.

2.10 MOTIVATION AND PROBLEM STATEMENT

The most important moral issue in the credit card trade is fraud involvement. The main aspires are, primarily, to recognize the different types of credit card fraud, and, secondly, to evaluate unconventional techniques that have been used in fraud detection. The sub-aim is to present, compare and examine recently published discovering in credit card fraud detection. Credit card fraud detection has developed a
number of techniques via bunch of investigate interest and, with special importance on, data mining and distributed data mining have been recommended.

Also there is a problem is, first the observation data possibly will be missing for a number of intervals. Following that there are multiple observation streams that are not necessarily synchronous to each other and possibly will have different “emission distributions” for the same state. So in proposed research we are using multiple observation sequences which are associated with the semi hidden state sequence and these observations may not be synchronized to each other.

We divide a large data set of labeled transactions (either fraudulent or legitimate) into smaller subsets by applying distributed data mining techniques to generate classifiers in parallel, and come together the resultant base models by meta-learning from the classifiers’ performance to produce a meta-classifier. In addition extensibility, combining multiple models computed over all available data produces meta-classifiers that can counterbalance the loss of predictive presentation that usually occurs when mining from data subsets or sampling. Furthermore, when we use the learned classifiers (for example, during transaction authorization), the base classifiers can carry out in parallel, with the meta-classifier then combining their results. So, our approach is highly efficient in generating these models and also relatively efficient in applying them.

In [53], author presented a survey of various techniques used in credit card fraud detection mechanisms and evaluates each methodology based on certain design criteria. The existing system of traditional detection method mainly depends on database system and the education of customers, which usually are delayed, inaccurate and not in-time. After that methods based on discriminate analysis and
regression analysis are widely used which can detect fraud by credit rate for cardholders and credit card transaction.

For a large amount of data it is not efficient. These resulted in high amount of losses due to fraud and the awareness of the relation between loss and the available limit which is need to be reduced. There is also necessity to reduce number of false alert in real time scenario. Thus they proposed system that overcomes the above mentioned issue in an efficient way. Using genetic algorithm the fraud is detected and the false alert is minimized and it produces an optimized result. The fraud is detected based on the customer’s behaviour. A new classification problem which has a variable misclassification cost is introduced. Here the genetic algorithms is made where a set of interval valued parameters are optimized. Their architecture diagram is given below:
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 no 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 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. 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 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. 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 fraud detection
problems. They consider naïve Bayesian, 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 FDS as a
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.

2.11 SURVEY ON EARLIER WORK

2.11.1 Hidden Markov Model (HMM)

HMM-based purposes are common in different areas such as speech recognition, bioinformatics, and genomics. In recent years, Joshi and Phoba [2] have explored the capabilities of HMM in abnormality detection. Cho and Park [3] proposed that improves the modeling time and performance which is an HMM-based intrusion detection system by considering only the privilege transition flows based on the domain knowledge of attacks. Ourston et al. [4] have suggested the application of HMM in identifying multistage network attacks. Hoang et al. [5] present a new method to process series of system calls for anomaly detection using HMM. Lane [6] has used HMM to model human behavior. Once human behavior is correctly modeled, any detected divergence is a cause for concern. Fujimaki and Morinaga [15] include recently a new Bayesian approximation inference method for fusion models. Through the factorized information criterion (FIC) and factorized asymptotic Bayesian inference (FAB). Each condition of the Markov chain gives rise to an release function of visible events (Rabiner and Juang, 1986)[16].

2.11.2 Semi-Hidden Markov Model (SHMM)

Hidden-Markov models are a version of HMMs capable of explicitly modeling the timing of state transitions is dealt in (Guedon, 2003) [17]. In cooperation HMMs and SHMMs have been shown to be competent of capturing time-varying signal characteristics by statistically sculpting the underlying active of the
signal (Rabiner, 1989) [16]. Importantly, HMMs and SHMMs are interpretable because (1) the emission function of each hidden state is expressed explicitly over the space of observable events and (2) the changes between hidden states are as well clearly modeled. By reviewing all the literature details there is a detection of accuracy is lacking. Jared O’Connell [21] the main features of the MHSMM package are developed as follows: Observations are allowed to be multi-variant. Missing values are allowed. Observations must be recorded at equidistant times. The package is designed to allow the specification of custom emission distributions. It is possible to have multiple sequences of data. [22] They believe that in most real applications we need models that combine the features of all of these models and introduce a new class of models, namely the hierarchical multichannel hidden semi-Markov models.

2.11.3 Expanded state Semi-Hidden Markov Model (ESHMM)

The multi-view approach [23] is based on the principle of maximizing the consensus among multiple independent hypotheses; they developed this principle into a semi supervised hidden Markov Perception algorithm. [24] Models that combine Markovian states with implicit geometric state occupancy distributions and semi-Markovian states with explicit state occupancy distributions are investigated. This type of model retains the flexibility of hidden semi-Markov chains for the modeling of short or medium size homogeneous zones along sequences but also enables the modeling of long zones with Markovian states. The forward–backward algorithm, which in particular enables to implement efficiently the E-step of the EM algorithm, and the Viterbi algorithm for the restoration of the most likely state sequence are derived. Lee-Min Lee [26] The HSMM is equivalent to a form of expanded state
HMM (ESHMM). Expansion of a state to several connected states with the same output probability distribution allows for more flexible duration model.

Yi Xie and Shun-Zheng Yu [27] an extended hidden semi-Markov model is proposed to describe the browsing behaviors of web surfers. In order to reduce the computational amount introduced by the model’s large state space, a novel forward algorithm is derived for the online implementation of the model based on the M-algorithm. In [33] a single repository data base where data is stored in central site, then applying data mining algorithms on these data base, patterns are extracted, which is clearly implausible and untenable for many realistic problems and databases. To deal with these complex systems has revealed opportunities to improve distributed data mining systems in a number of ways. Furthermore, in certain situations, data may be inherently distributed and cannot be merged into a single database for a variety of reasons including security, fault tolerance, legal constraints, competitive reasons, etc. In such cases, it may not be possible to examine all of the data at a central processing site to compute a single global model. Here, we develop techniques that scale up to large and possibly physically distributed databases.

2.11.4 Neural networks

A lot of research notice and a number of techniques have developed to the Credit card fraud detection, with unique prominence on data mining and neural networks, have been proposed. Recently, Syeda et al [39] have employed parallel granular neural networks (PGNNs) for civilizing the rapidity of data mining and knowledge discovery process in credit card fraud detection. Aleskerov et al [42] developed CARDWATCH, a database mining system employed for credit card fraud
detection which is based on a neural learning module, offers a crossing point to the variety of commercial databases. Ghosh and Reilly [37] have proposed credit card fraud detection with a neural network which is trained on a large illustration of labeled credit card account transactions. These transactions enclose instance fraud cases due to gone cards, stolen cards, purpose fraud, forged fraud, mail-order fraud and non-received issue fraud. Where a drawback is that a complete system has been implemented for this purpose which is time consuming. Stolfo et al [40] recommend a credit card fraud detection system (FDS) by means of meta-learning techniques to discover models of fraudulent credit card transactions.

A general strategy that provides a means for combining and integrating a number of separately built classifiers or models is called as Metalearning. This will be trained correlation of the predictions of the base classifier. The equivalent collection has also labored for fraud and intrusion detection on a cost-based model [41]. Kim and Kim [43] have recognized twisted distribution of data and mix of legal and fraudulent transactions as the two main motivations for the complexity of credit card fraud detection. Supported on this observation, they utilize fraud density of real transaction data as a confidence value and produce the weighted fraud score to diminish the number of misdetections. Fan et al [44] suggested the application of distributed data mining in credit card fraud detection. Brause et al [45] have extended an approach that entails advanced data mining techniques and neural network algorithms to attain high fraud coverage. Chiu and Tsai [46] have proposed web services and data mining techniques to initiate a collaborative scheme for fraud detection in the banking industry. By means of this system, participating banks partition knowledge about the fraud patterns in a heterogeneous and distributed environment.
Phua et al [47] have done a research based a widespread survey of existing data mining supported fraud detection systems and published an inclusive information. Prodromidis and Stolfo [47] exploited an agent based move toward with distributed learning for detecting frauds in credit card transactions. For achieving higher accuracy it is supported on artificial intelligence and inductive learning algorithms and Meta learning methods is combined. Phua et al [51] proposed the use of Meta classifier similar to [40] in fraud detection troubles. They believe naïve Bayesian, C4.5 and Back Propagation neural networks as the base classifiers.

Vatsa et al [52] have recently planned a game theoretic approach to credit card fraud detection. They model the communication among an attacker and a fraud detection system as a multi-stage game between two players, each trying to take full advantage of his bribe. The difficulty with most of the above-mentioned approaches is that they want labeled data for both authentic as well as fraudulent transactions to train the classifiers. Receiving real world fraud statistics is one of the major harms connected with credit card fraud detection. In addition, these approaches cannot discover new categories of frauds for which branded data is not accessible.

Credit card fraud detection has drawn a lot of research interest and a number of techniques, with special emphasis on data mining and neural networks, have been suggested. 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 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. 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 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. 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 et al. Suggest the use of meta classifier similar to in fraud detection problems. They consider naïve Bayesian, 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 FDS as a multistage 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. We model 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. Otherwise, due to the “base rate fallacy” effect, bank administrators may tend to ignore the alarms. To the best of our knowledge, there is no other published literature on the application of HMM for credit card fraud detection.

2.12 SUMMARY

In this chapter of literature review, we analyzed several papers that are formerly involved in credit card fraud detection. However multiple observation streams of these systems do not have necessity of being synchronous to each other and lack in effective prediction of fraud detection. This detecting efficiency is improved by using three variants of Hidden Markov Model which are mentioned in upcoming chapter.