CHAPTER 9

CONCLUSION

We have proposed an application of different Hidden Markov Model 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, while 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. We implemented this credit card fraud detection with three variants of HMM that are listed below:

- Semi Hidden Markov Model (SHMM)
- Multiple Semi Hidden Markov Model (MSHMM)
- Advanced Hidden Markov Model (AHMM)

Hidden semi-Markov model is introduced into credit card fraud detection systems. We present an algorithm of fraud detection based on HSMM, which computes the distance between the processes monitored by intrusion detection system and the perfect normal processes. In this algorithm, based on maximum entropy principle (MEP), we introduce the concept of average information entropy (AIE), which is used as detection metric via analyzing variable-length observed symbol sequences. To improve accuracy, we propose a new approximation inference
algorithm and refer to it as a factorized asymptotic Bayesian inference (FAB) with the HSMM.

The fundamental assumption in the existing HMM and SHMM models is that there is at least one observation produced per state visit and that observations are exactly the outputs of states. In some applications, these assumptions are too restrictive. We extended the ordinary HMM and SHMM to the model with missing data and multiple observation sequences. In our project Multiple Semi Hidden Markov Model is implemented to provide better fraud detection mechanism. Where the multiple observations will be collected using distributed data mining technique and the detection phase is successfully executed.

From the calculation of the probability values the training value and the testing value will be compared. If two values are same then there is no anomaly detection other than that there can be a false alarm generation to denote the anomaly involvement. Considering existing research they modeled the sequence of operations in credit card transaction processing using a Hidden Markov Model (HMM) and shown how it can be used for the detection of frauds.

To provide better accuracy and to avoid computational complexity in fraud detection in proposed work semi Hidden Markov model (SHMM) algorithm of anomaly detection is presented which computes the distance between the processes monitored by credit card detection system and the perfect normal processes. There is also a larger class of practical problems when they are modeled in the setting of SHMM. Also the major drawback is that it generally assumes that there survives at least one observation connected with every state that the hidden Markov chain takes on.
The credit card transaction method is examined as the basic stochastic process of an (Advanced Hidden Markov Model) AHMM. The variety of transaction quantity considered as the observation symbols, while the kinds of item have been deemed to be states of the AHMM. In addition to comprise recommended a technique for decision the spending profile of cardholders is authorized or not. As well as purpose of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters with the best fit observation is that providing an effective credit card fraud detection system.

It has also been enlightened how the HMM vary with the AHMM can detect whether an arriving transaction is fake or not. Experimental results show the performance and effectiveness of AHMM system and show the efficiency of knowledge the spending profile of the cardholder in AHMM system. To improve the efficiency of SHMM then we combined the multiple observation of SHMM called Multiple Semi Hidden Markov Model (MSHMM) through this we can improve the detection accuracy better than the SHMM. In SHMM there is a problem that it cannot find an optimal state sequence for the underlying Markov process also this observed sequence cannot be viewed as training a model to best fit the observed data.

Then, we aimed to model the sequence of observations in credit card transaction processing using an Advanced Hidden Markov Model (AHMM) and which can be utilized for the exposure of frauds. In this process an AHMM is initially trained with the regular manners of a cardholder. If an incoming credit card transaction is not recognized by the trained AHMM with adequately high probability, it is believed to be fraudulent. This work desire to regulate the model parameters to best fit the observations.
Finally, the model parameters are chosen in the optimized manner for the efficient prediction of credit card fraudulent behavior. The optimized selection is done by introducing the methodology called the cuckoo search optimization algorithm which aims to satisfy the constraint of the server in terms of the credit card fraudulent behavior. Optimized Multiple Semi-hidden Markov model is proposed to optimize the model parameters so that very effectively detect the anomaly. Because it is significant to optimize the model parameters so that to improve the accuracy in the detection performance. The MSHMM model parameters are decided during an iterative process which is called training phase. A cuckoo search optimization technique is utilized to optimize the Multiple Semi-hidden Markov model parameters.

The ranges of the matrices \( N \) and \( M \) are fixed but the elements of \( A, B \) and \( \pi \) are to be decided, focus to the rank stochastic condition. The information that can efficiently re-estimate the model itself is one of the more incredible features. 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% over a wide variation in the input data. The system is also scalable for handling large volumes of transactions.