IJICT International Journal of Information & Communication Technology

ITRC

Volume 2- Number 3- November 2010 (pp.21-28)

A Timelier Credit Card Fraud Detection by Mining Transaction Time Series

Leila Seyedhossein
School of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
l.seyedhossein@ut.ac.ir

Mahmoud Reza Hashemi
School of Electrical and Computer Engineering
University of Tehran
Tehran, Iran
rhashemi@ut.ac.ir

Received: September 18, 2010 – Accepted: November 20, 2010

Abstract— **As e-commerce sales continue to grow, the associated online fraud remains an attractive source of revenue for fraudsters. These fraudulent activities impose a considerable financial loss to merchants, making online fraud detection a necessity. The problem of fraud detection is concerned with not only capturing the fraudulent activities, but also capturing them as quickly as possible. This timeliness is crucial to decrease financial losses. In this research, a profiling method has been proposed for credit card fraud detection. The focus is on fraud cases which cannot be detected at the transaction level. Based on the fact that there are strong periodic patterns in cardholders' behavior, the time series of aggregated daily amounts spent on an individual credit card has been considered in the proposed method. In this method, the inherent periodic and seasonal patterns are extracted from the time series to construct a cardholder's profile. These patterns have been used to shorten the time between when a fraud occurs and when it is finally detected. Simulation results indicate that the new approach has resulted in a timelier fraud detection, improved detection rate and consequently less financial loss in the cases where a cardholder follows a regular or semi regular periodic behavior. The proposed method is equally applicable to other e-payment methods with minor application-specific modifications.**

Keywords- Fraud detection; aggregatioal profile; time series;

I. Introduction

In recent years, fraud detection has always been a hot topic in the context of electronic payments. This is mostly due to considerable financial losses incurred by payment card companies as a result of fraudulent activities. According to a CyberSource study conducted in 2011, the amount of loss incurred due to payment fraud in the United States and Canada has been $2.7 billion in 2010, which is a considerable amount [1].The growing value of fraudulent activities is partly due to the increasing volume of online shopping by consumers, making it an inviting source of revenue for criminals. Also modern technologies make fraudsters more sophisticated and more difficult to be detected.

The focus of this research is on credit card payments as one of the widely accepted electronic payment methods. Based on a survey conducted by the Consumer Payments Research Center of the Federal Reserve Bank of Boston, 72.8 percent of American consumers have used a credit card in 2008 [2]. Credit card fraud constitutes of using the credit card of a valid cardholder for personal benefit without the owner's knowledge and consent. The fraudulent activities affects all parties in the credit card payment cycle; namely cardholder, merchant, issuer and acquirer. Among them the merchant is the most vulnerable party [3]. Credit card frauds can be divided broadly in the following categories: lost/stolen cards, account takeover, card not present (CNP), fake and counterfeit cards, mail not-received and application frauds. Card-not-present is the most common type, where the merchant no is longer protected by the advantages of physical verification [4].

A good fraud detection system should be able to identify the above mentioned assortment of fraudulent activities accurately, and more importantly as quickly as possible. Fraud detection approaches can be divided into two main categories: misuse detection and anomaly detection [5]. A misuse detection system is trained on examples of normal and fraudulent transactions. So they can only recognize known frauds. While an anomaly detection system is trained only on normal transactions and they have a potential to detect novel frauds. Difficult access to labeled data and the evolving nature of fraudulent activities, has led to more concentration on anomaly detection techniques. In these techniques the cardholder's profile is constructed based on his normal spending habits and any inconsistency with regards to this normal profile is considered as a potential fraud. The problem with this approach is the large number of false alarms due to normal changes in cardholder's behavior.

Using anomaly detection techniques for fraud detection involves constructing an efficient profile which considers all aspects of a cardholder behavior. Usually a fraudster is not familiar with the spending habits of a cardholder when it tries to gain the most profit from a stolen card. Hence they tend to perform high value transactions, which usually have a different characteristic from the normal cardholder transactions. In this context the transactional profile can reveal the frauds. Many researches have considered this kind of fraudulent activities and constructed a transactional profile [6], [7], [8] and [9]. But more cautious fraudsters try to follow the normal behaviors of cardholders or perform low value transactions in short time intervals. In this case the frequency or volume of transactions is a much better indicator of fraud compared to the characteristics of each individual transaction. For instance, in these frauds the total number or total amount spent on a credit card over a specific time window increases. A few researches consider this type of frauds and construct an aggregated profile [4] and [10]. In these schemes the two approaches are combined together to improve detection rate and accuracy.

Clearly, a fraudster wishes to gain the most profit from a stolen card. Hence, he continues to commit fraud until it is being caught and the card is cancelled. Consequently, a compromised account incurs more loss as time passes. Hence, an efficient fraud detection system should raise a fraud alarm as early as possible [11]. Otherwise, a significant amount of money may be stolen. The faster the fraud detection system is the less significant the financial ramifications are. That is why the response time of a fraud detection system, referred to as timeliness is one of its most crucial characteristics.

Most aggregation approaches have a relatively slow detection, mostly due to their data collection latency. This problem seems more significant when the aggregation period is longer. This implies that a method with a shorter aggregation period is desirable. But smaller aggregation may reduce detection accuracy by not incorporating adequate information in the detection process.

Furthermore, using an aggregation technique involves discarding some useful information, such as the order of transactions. The transaction order may be used to identify a pattern in fraudulent activities.

In this research, we apply the order of transactions to shorten the aggregation period, which leads to a timelier fraud detection. For this purpose the sequence of aggregated daily amounts spent on an individual cardholder in a time window has been considered. Then the inherent patterns in these time series have been extracted to shorten the time between when a fraud occurs and when it is finally detected. We have taken advantage of the order of data for timelier fraud detection. We demonstrate that the proposed approach leads to improved detection rate, and timeliness while it decreases the cost involved in some circumstances.

Although in this paper we have focused on applying the proposed method on credit card fraud detection, but we believe that the proposed approach can be used for fraud detection in other e-payment mechanisms with minimal application specific modifications. Similarly, the same scheme is applicable for fraud detection in other applications where there is an inherent periodic, seasonal or geographic pattern in the consumer behavior, such as insurance and telecommunication.

The rest of the paper is organized as follows. Section 2 presents related works on credit card fraud detection. The proposed system has been introduced in section 3. Section 4 illustrates the experimental results. Finally, the paper concludes with future works and conclusion remarks in sections 5.

II. Related Works

Misuse detection and anomaly detection are the two main approaches used for credit card fraud detection. The emphasis on misuse detection approaches is usually upon applying classification methods at the transaction level. For a recent survey of applying misuse detection techniques in the area of credit card fraud detection one may refer to [12], [13], [14], [15], [16] and [17]. In these researches various classification methods such as neural networks, decision trees, logistic regression and support vector machine have been used and compared against each other in the area of credit card fraud detection. Also in a recent research in [11] various classification methods have been applied on aggregated transactions. This research has demonstrated that aggregated values are a better indicator of frauds in some circumstances.

Among the researches which have been conducted on credit card fraud detection we have concentrated on the ones which apply anomaly detection techniques, the so called behavioral or profile-base techniques. Typically they have constructed a cardholder profile based on normal training data and then tried to detect fraudulent activities based on the inconsistencies with the normal behavior. A recent publication in [18] has compared the performance of a misuse detection approach with the performance of a profile-base approach and concluded that the latter is more suitable for large scale and evolving nature of plastic card fraud detection.

Most profile-base approaches have applied data mining techniques like clustering and association rules to construct a transactional profile. For instance, in [6] self-organization map has been used to cluster customer transactions. The density of each cluster is the basis of distinction between normal and rare behavior of customers which can be used to detect suspicious activities. The emphasis of this research is on real-time fraud detection to reach a cost effective system. A similar approach has been applied in [19] to create a model of typical cardholder behavior and to detect suspicious transactions based on the deviation from this normal behavior. Also in [7] DBSCAN, which is a density based clustering algorithm, has been used to create clusters of customer transactions and build a transactional profile. An example of using association rules can be found in [8]. In this research, recent transactions of a customer have been dynamically profiled using association rules, to indicate how unusual a new transaction is. The word recent is defined by a sliding window. Also in [20] fuzzy association rules have been applied to extract cardholder behavior patterns for credit card fraud detection.

In few researches in this area, the sequence of transactions has been considered for building customer profiles. An example of which can be found in [9]. In this research a Hidden Markov Model for each customer has been built during the training phase based on a sequence of transaction amounts. When a new transaction arrives, a new sequence is constructed by dropping the first member of the old sequence and appending the new transaction at the end. If the new sequence is not accepted by the trained model, it is considered as fraud. In another research in [5] which combines anomaly and misuse detection techniques, normal and fraudulent sequences of quantized transaction amounts have been formed to capture the cardholder behavior. Then a sequence alignment technique has been used to measure the similarity between a new sequence and the training model. In a different research in [21] for each target cardholder, sequences of daily transaction amounts have been compared against the other cardholders to find the k nearest ones. These similar sequences have been grouped to form the peer group of that cardholder. If the future sequences of that cardholder deviate from its peer group, a fraud alarm is raised. The basis of this research is the assumption that when a group of cardholders are behaving similarly until a specific time, it is very likely that they will continue to have the same behavior for a while.

A few researches construct an aggregational profile. For instance in a hybrid system proposed in [4], the aggregated profile is built based on the average spending amount of the cardholder in all k-day rolling window of his transaction history. The superiority of this approach over a fixed window lies on the timeliness of detection which is mentioned to be crucial for a fraud detection system. Also in [10] a hybrid profile has been proposed which consists of transactional and aggregational parts. In this research the aggregational profile has been constructed based on total weekly count and amount of cardholder transactions.

To the best of our knowledge, among the researches which apply aggregational profile, none of them has considered the sequence of data. This has been addressed in this research.

III. PROPOSED METHOD

In this research, we have explored the application of transaction sequence for the purpose of timelier credit card fraud detection. The focus in this work is on fraud cases which cannot be detected at the transaction level. In fact, we have proposed an improved aggregated profile which exploits the inherent patterns in time series of transactions. Some extensive modeling on real data reveals strong weekly and monthly periodic pattern in cardholder spending behavior [22]. Based on these observations we believe that instead of looking at individual transactions, it makes more sense to look at sequences of transactions. But it is impractical to consider the entire series of cardholder transactions because of the high dimensionality of this data. So we model the time with a sequence of aggregated transactions which can reduce the dimensionality. Also aggregated transactions are more robust to minor shift in cardholder behavior.

To form the time series, the total amount of transactions in each day of year has been calculated. Then the ordered series of these aggregated values form the time series. Like the aforementioned researches in [4] and [10] which consider 7 days for aggregation, we form 7-day time series. So each time series consists of 7 dimensions each of which corresponds to the total amount of transactions in one day.

As mentioned before, based on observations on real data, there are some periodic structures in transactions, so we expect to find similar trends in 7-day time series year after year. Also since the first year of each year is considered as the starting point of the 7-day period of that year, the time series for each year would be different in terms of days of the week. For example one year may start on Sunday while the next year starts on Friday. This implies that for each year the 7-day time series, of a cardholder that follows a stable weekly trend, should be aligned in terms of days of the week accordingly. Furthermore, a cardholder himself may have some shift in purchasing days. Another pattern is some occasional behavior that can be seen due to holidays and occasions which are repeated in all years like the New Year holidays [22].

In this research, the objective is to extract these inherent patterns in time series of aggregated transactions, and apply them to detect fraudulent activities faster and more accurately. In fact, by exploiting these patterns we can detect fraud cases before the end of an aggregation period. The details of constructing profiles and the fraud detection method will be explained in the following sub sections.

A. Profile Extraction

To construct a cardholder profile, his normal transactions are being considered as training data. As mentioned earlier a preprocessing step is performed to build the corresponding time series. Then the inherent patterns in these time series should be extracted to build an efficient profile. In this research, two possible patterns are extracted from the training data in two steps. The first possible inherent pattern in a 7-day period could be following the same trend in all years. For extracting this pattern, time series have been clustered using k-means [23], one of the most popular clustering algorithms, with the Euclidean distance. Since the Euclidean distance is used as the similarity measure, the time series which have almost the same trend will be placed in the same cluster. After clustering, if all yearly time series for a specific 7-day period are placed in the same cluster, this period is labeled as a stable-trend period. Then all of the time series that belong to these periods are excluded from the training data and the other ones remain for further analysis in the next stage.

The Euclidean distance is very sensitive to small distortions in the time axis. If two time series are identical, but one is slightly different along the time axis, then the Euclidean distance may consider them to be very different from each other. But as it was mentioned before, the second possible inherent pattern in a period could be following the same trend by permuting the time axis as we can see in Fig.1. In order to find the similarity between such sequences, the time axis should be best aligned before calculating the Euclidean distance.

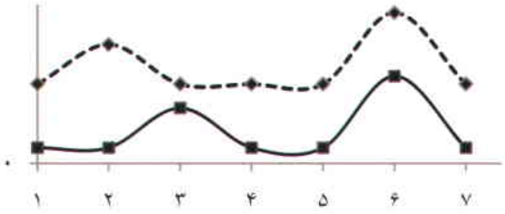

Figure 1. An example of permuted-trend time series

The remaining time series from the first stage have been clustered using this new distance, referred to as permuted distance. Typically, the k-means algorithm selects k initial points as cluster centers. Then each point is assigned to the closest center using a distance measure. When all points have been assigned, the new centers are recalculated by averaging cluster members. These steps are repeated until the centers no longer move. Usually the Euclidean distance is used as the distance measure in the k-means algorithm. This should be modified for the permuted pattern. To find the permuted distance between two time series, any permutation of the time axis for the first one is considered, and the Euclidean distance between all of them with the second one is calculated. Then the minimum value is selected as the distance between the two time series. Also the current averaging method for finding new centers may not produce the real average of the time series in our case, thus resulting in incorrect k-means clustering results. See Fig. 2 and Fig. 3 for further clarification. Fig. 2 is the result of usual averaging method of the two time series while we expect the result which is shown in Fig. 3. So the time series should be aligned along the time axis before calculating the average time series.

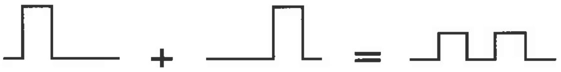

Figure 2. Conventional averaging of the two time series

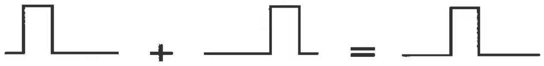

Figure 3. Desired averaging of the two time series

The remaining time series from the first stage are clustered with this modified k-means. As a result the time series which are almost the same after alignment along the time axis are placed in the same cluster. We labeled the 7-day periods for which all yearly time series are placed in the same cluster as permuted-trend. Moreover there are some yearly occasions in which the cardholder behavior is almost the same for all years like the New Year holidays. The cardholder profile can be further improved by identifying these occasions in permuted-trend periods. For this purpose, when the time series of a period are permuted to compute the permuted distance, the best alignments are considered. The stable days are the ones which are not permuted in the best alignments.

After these two stages, the remaining periods are labeled as unpredictable-trend. Hence, at the end of the training phase we have a time series for each 7-day period of the year with the specification about which trend it belongs to and which days are stable days for the second trend.

The mining method presented in this section is applicable in other sequence-related data mining applications where it is required to extract a periodic behavior. Also the profiling method can be applied in any application that demonstrates an inherent periodic behavior.

B. Fraud Detection

After the training phase, fraudulent activities can be detected based on the degree of deviation from the cardholder profile. For this purpose when new transactions arrive they are accumulated to build the current period's time series. Based on the type of current period in the cardholder's profile, which can be stable-trend, permuted-trend or unpredictable-trend, the fraud detection is performed in real time at the end of each day or at the end of the period respectively. For the stable-trend periods, since the cardholder behavior in corresponding days are almost the same, the fraud detection can be performed in real time. As the transactions of a day are accumulated, they can be compared against the corresponding values in the profile. Whenever this value exceeds by a factor of θ_1 from the corresponding amount in the cardholder's profile, it indicates a fraud. For the permuted-trend periods, at the end of each day, the similarity between the current time series with the corresponding one in the profile is computed. Since in the middle of a period the current time series is smaller than the corresponding one in the profile, we should consider all of the subsets of profile time series with the same length as the current time series. Then the minimum permuted distance between them is considered. If this value exceeds a threshold θ_2, it indicates a fraud. While considering all subsets of profile time series, the days which are flagged as stable-days should remain immovable. As a result of the above approach, at the end of each day we can determine whether or not a fraud has occurred up to this point and there is no need to wait until the end of a period. One important point is that for this trend when we are building the time series, whenever a fraud case has been identified in a day, we should replace this day with the corresponding value from the profile in order to prevent the fraud value from affecting the decision for the upcoming days of the period.

Finally, for the unpredictable-trend periods at the end of a 7-day period, we compute the distance between the current time series and the corresponding one in the cardholder profile and if it exceeds a threshold θ_3, it indicates fraud. For this group, at the end of the period we have a label which tells us there are some frauds in this period.

The best value for the parameters θ_1, θ_2 and θ_3 is obtained by examining the performance of the system over a range of threshold values using a tuning data set and selecting the values that generate the best average results.

Clearly, the proposed method improves the timeliness of fraud detection which is proved to be most effective in the stable-trend periods and the permuted-trend, consecutively. In these two cases, the system does not have to wait until the end of the aggregation period to detect fraudulent activities and it can tackle frauds in real time or at the end of each day without a considerable increase in false alarms. This is mostly due to considering the order of transactions as another cardholder's behavioral pattern in addition to his aggregated spending pattern. However, the mentioned method does not improve the timeliness of fraud detection for the unpredictable-trend periods.

IV. EXPERIMENTAL RESULT

The performance of the proposed scheme has been compared with the results of [4] and [10] since they have both used an aggregational profile.

In [10] the aggregated profile is constructed based on the total weekly count and amount of each cardholder's transactions. Then at the end of each week the current aggregated value is updated and compared with the profile value to detect considerable deviations which can be considered as a potential fraud. We expect that our proposed method should increase the detection rate and improve the timeliness of the method in [10] in the situation where the cardholder follows a periodic pattern. Also the aggregational profile proposed in [4] will be compared against our proposed method. In [4] the model of aggregation consists of a set of descriptors for quantifying time series of cardholder behavior. These time series are built using all of the k-day periods of normal transactions. 1, 3 and 7 days periods are used for evaluation, among them we choose the 7-day period for comparison, which conforms to our approach.

A. Dataset

To evaluate our work we have developed an application to generate synthetic data containing genuine and fraudulent transactions. The profile driven method has been used for generating data like the one applied in [8]. We believe that our dataset provides a good approximation for evaluation of the proposed method because we use real scenarios to create the data. As it was mentioned before, based on observations from real data, there are some periodic structure in credit card transaction data and also some occasional events [22]. Also there are various weekly and seasonal patterns in cardholder behaviors [24]. These real scenarios have been implemented in data generation to justify the results. Also normal distribution, which is the most common observed probability distribution in many natural processes, has been used to create number and amount of transactions.

Five attributes for each transaction have been considered including year, month, week of month, day of week and amount. The first four attributes indicate the time sequence of data and the last one is a good descriptor to quantify the time series. We have created four different profiles to generate different kinds of cardholder behaviors. In the first one the cardholder has almost similar periodic behavior. In the second one the cardholder has similar behavior with some shift in the time axis. In the third profile the cardholder has an unpredictable behavior. Finally, the fourth cardholder has a mixture of different behaviors in different times. Transactions for three years are created for each cardholder as training data. Then a mixture of genuine and fraudulent transactions of one year is generated for test data. Fraudsters usually follow two different scenarios to avoid detection: high value transactions with long gaps or small value ones with short gaps. The first scenario can be detected by a transactional profile and the second one can be detected by an aggregational profile. Because we want to evaluate an aggregational profile, fraudulent activities are created based on the second scenario.

For each profile three datasets are created. The first one which contains normal transactions is a training set. The second and third ones contain a mixture of normal and fraudulent transactions. The second data set is used for tuning and obtaining the best values for the system parameters. The third data set is a test set used for evaluating the proposed method. Table 1 shows the number of transactions in each dataset of the four profiles.

TABLE I. CHARACTERISTICS OF THE DATASET

	Profile 1	*Profile 2*	*Profile 3*	*Profile 4*
Training Set	3314	5299	6709	3951
Tuning Set	1186	2129	2950	1854
Test Set	1206	2114	2968	1809

B. Performance Measures

The transactions which are flagged by a fraud detection system include fraudulent and normal transactions which are classified correctly (TP, TN) and fraudulent and normal transaction flagged erroneously (FN, FP). A good fraud detection system should lead to maximum number of TP and TN and minimum number of FP and FN.

Several performance measures have been applied for fraud detection systems. The appropriate one should take into account the specific issues in fraud detection systems. In a recent research [25] the appropriate performance measures for plastic card fraud detection systems have been proposed. We have applied two measures which are proposed in that research and widely applied in recent fraud detection researches: timeliness ratio and loss function. The first one measures the speed of fraud detection and defined as the proportion of FN to F. the second one measures the cost involved. In this measure different costs are considered for different errors. Generally, FNs are more serious than FPs. We use the function used in [6] and presented here in (1).

$$L(s) = \frac{TP + FP + 100 * FN}{N + 100 * F} \qquad (1)$$

A smaller value for these two measures indicates a better performance. Also we use a standard measure, TP%, which is the percent of TPs to all of the fraudulent transactions. A higher value indicates a better performance.

C. Optimization of Parameters

As discussed in section III, the proposed method has 3 parameters, θ_1, θ_2 and θ_3. In choosing a value for these parameters, there is a tradeoff between TP% and FP%. In this work we choose the best value for each of them using TP/FP(%). The best value for each parameter is obtained by examining the performance of the system over various values of them using the tuning sets and choosing the one with the best average result on all of the profiles. As a result the values 1.4, 0.7 and 0.2 have been obtained experimentally for θ_1, θ_2 and θ_3, respectively.

D. Validation Results

First we study the performance of our aggregation method to the one proposed in [10]. As we can see in Fig. 4 TP% of the proposed method is better than the one proposed in [10]. Also Fig. 5 and 6 indicate that the cost and timeliness of our proposed method is better too. It can be clearly seen from these figures that when a cardholder follows an almost stable trend in the corresponding times of the year, the case which has occurred in the first profile, the performance of the system increases significantly. This is mostly due to the fact that the fraudulent activities can be detected in real time. As a result, more frauds can be detected by the system, in a timelier manner and with less cost. In the second test case which indicates a cardholder with the permuted behavior, the performance of the system is slightly better, because the fraudulent activities can be detected at the end of each day. But if the cardholder has an unpredictable behavior, which is simulated in the third case, the performance of our method is almost the same as the one proposed in [10] because there is no pattern in the cardholder behavior which can be used for timelier detection and the fraudulent activities can be detected at the end of 7-day periods.

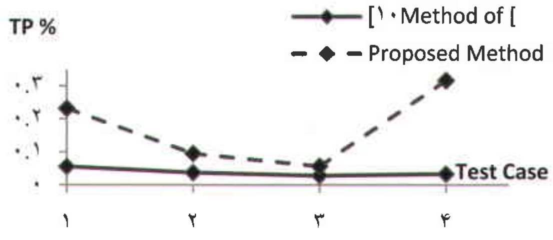

Figure 4. TP% of four test cases for aggregational part of the method proposed in [10] against our method

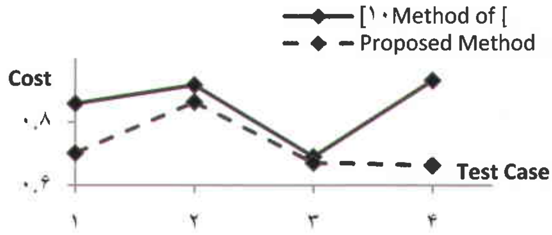

Figure 5. Cost of four test cases for aggregational part of the method proposed in [10] against our method

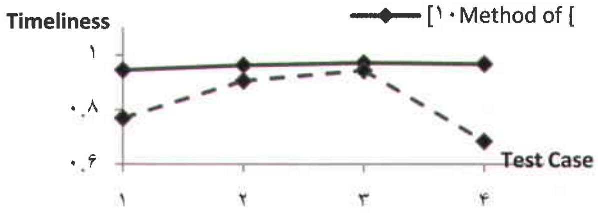

Figure 6. Timeliness of four test cases for aggregational part of the method proposed in [10] against our method

Next the performance of the proposed method is compared against the aggregation method proposed in [4]. In that research the procedure for detecting fraudulent activities is run at the end of each day, considering 7 days before the current day. It can be seen from Fig. 7, 8 and 9 that almost the same results are obtained as the previous experiment. One of the underlying reasons for this improved result may be due to the consideration of seasonal behaviors in the proposed method. Also the same reasons as discussed for the previous experiment apply to this experiment as well.

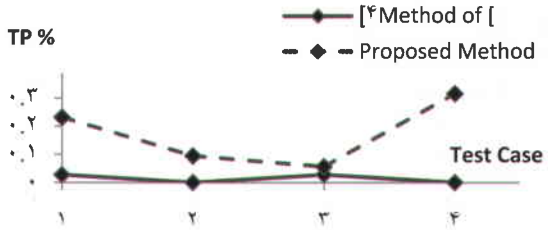

Figure 7. TP% of four test cases for aggregational part of the method proposed in [4] against our method

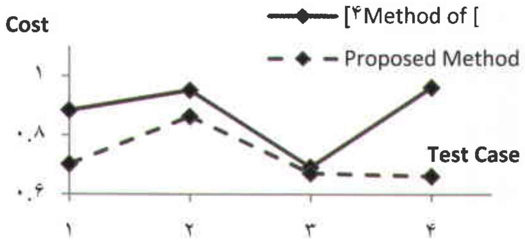

Figure 8. Cost of four test cases for aggregational part of the method proposed in [4] against our method

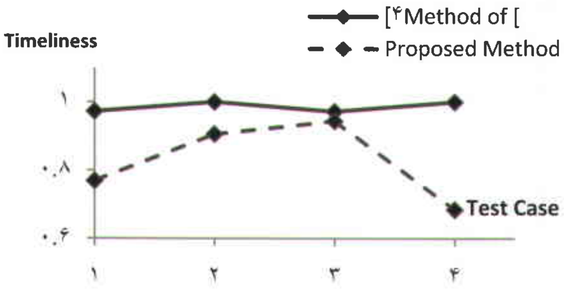

Figure 9. Timeliness of four test cases for aggregational part of method proposed in [4] against our method

V. Conclusion and Future Work

In this paper we have addressed the general problem of credit card fraud detection using anomaly detection techniques, by exploiting the sequence of transactions in constructing cardholders' profiles. We have investigated how this affects detection performance. The focus is on fraud cases which cannot be detected at the transaction level. A new method for constructing an aggregated profile is also proposed. To this end the pattern of aggregated daily purchases of cardholders are extracted from the training data. Due to the seasonal behavior of cardholders these patterns are time dependent. Then these extracted patterns have been used for more accurate fraud detection in a timelier manner. Experimental results show that the proposed method can improve the fraud detection in the situations where cardholders follow some purchasing patterns in corresponding times of the years.

Cardholder behavior changes over time. This can be due to changes in their life style or other environmental factors. Hence, one way to improve the current work is by detecting or even predicting these behavior changes ahead of time and updating the profiles accordingly. Furthermore, in this paper we have considered two behavior patterns only. The performance can be improved by introducing and detecting more patterns in customer spending.

In addition to a cardholder there are other stakeholders in a payment system. By incorporating these stakeholders in the fraud detection system we can improve the detection rate and even its timeliness. Hence, the current research can be extended by extracting the profile of these other entities, and considering them in the detection process.

Acknowledgment

The authors would like to express their gratitude to Iran Telecommunication Research Center (ITRC) for their support during the course of this research.

References

[1] CyberSource;"12th Annual Online Fraud Report"; 2011. http://forms.cybersource.com/forms/FraudReport2011NACYBSwww2011

[2] Kevin Foster, Erik Meijer, Scott Schuh, and Michael A. Zabek,"The 2008 Survey of Consumer Payment Choice", Federal Reserve Bank of Boston, 2010. http://www.bos.frb.org/economic/ppdp/2009/ppdp0910.htm

[3] J.Dara, L.Gundemoni, "Credit card security and e-payment: enquiry into credit card fraud in e-payment", Master's Thesis, Dept. Business Administration and Social Sciences, Lulea University, Sweden, 2006.

[4] M. Krivko, "A hybrid model for plastic card fraud detection systems," Expert Systems with Applications, vol. 37, no. 8, 2010, pp 6070-6076.

[5] A. Kundu, S. Sural, and A. Majumdar, "Two-stage credit card fraud detection using sequence alignment," Information Systems Security, Springer Berlin / Heidelberg, 2006, pp. 260-275.

[6] J. Quah and M. Sriganesh, "Real-time credit card fraud detection using computational intelligence," Expert Systems with Applications, vol. 35, no. 4, 2008, pp. 1721-1732.

[7] S. Panigrahi, A. Kundu, S. Sural, and A. Majumdar, "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning," Information Fusion, vol. 10, no. 4, 2009, p. 9.

[8] J. Xu, A.H. Sung, and Q. Liu, "Behaviour mining for fraud detection," Journal of Research and Practice in Information Technology, vol. 39, no. 1, 2007, pp. 3-18.

[9] A. Srivastava, A. Kundu, S. Sural, and A. Majumdar, "Credit card fraud detection using Hidden Markov Model," IEEE Transactions on Dependable and Secure Computing, vol. 5, no. 1, 2008, pp. 37-48.

[10] L.Seyedhossein, M.R. Hashemi, " A hybrid profiling method to detect heterogeneous credit card frauds", 7th International ISC Conference on Information Security and Cryptology, 2010, pp 25-32.

[11] C. Whitrow, D.J. Hand, P. Juszczak, D. Weston, and N.M. Adams, "Transaction aggregation as a strategy for credit card fraud detection," Data Mining and Knowledge Discovery, vol. 18, no. 1, 2009, pp. 30-55.

[12] R. Brause, L. T., and M. Hepp, "Neural data mining for credit card fraud detection," 11th IEEE International Conference on Machine Learning and Cybernetics , vol. 7, 2008, pp.3630-3634.

[13] R. Chen, S. Luol, X. Liang, and V.C. Lee, "Personalized approach based on SVM and ANN for detecting credit card fraud", International Conference on Neural Networks and Brain, 2005, pp. 810-815.

[14] A. Shen, R. Tong, and Y. Deng, "Application of classification models on credit card fraud detection," International Conference on Service Systems and Service Management, June 2007, pp. 1-4.

[15] M.F. Gadi, X. Wang, and A.P. Lago, "Comparison with parametric optimization in credit card fraud detection," Seventh International Conference on Machine Learning and Applications, 2008, pp. 279-285.

[16] M. Syeda, Z. Yan-Q, and P. Yi, "Parallel granular neural networks for fast credit card fraud detection," Proceedings of the IEEE International Conference on Fuzzy Systems, 2002, pp. 572 – 577.

[17] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J.C. Westland, "Data mining for credit card fraud: A comparative study," Decision Support Systems, vol. 50, no. 7, 2011, pp. 602-613.

[18] N.M. Adams, D.J. Hand, C. Whitrow, and D.J. Weston, "Off-the-peg and bespoke classifiers for fraud detection," Computational Statistics & Data Analysis, vol. 52, no. 9, 2008, pp. 4521-4532.

[19] V. Zaslavsky and A. Strizhak, "Credit card fraud detection using self- organizing maps," Information and Security, vol. 18, 2006, pp. 48-63.

[20] D. Sanchez, M. Vila, L. Cerda, and J. Serrano, "Association rules applied to credit card fraud detection," Expert Systems with Applications, vol. 36, no. 2, 2009, pp. 3630-3640.

[21] D.J. Weston, D.J. Hand, N.M. Adams, C. Whitrow, and P. Juszczak, "Plastic card fraud detection using peer group analysis," Advances in Data Analysis and Classification, vol. 2, no. 1, 2008, pp. 45-62.

[22] Niall M. Adams, David J. Hand, Giovanni Montana, David J. Weston, "Fraud detection in consumer credit", Expert Update SGAI, Special Issue on the 2nd UK KDD Workshop; vol. 2, no. 1, 2006.

[23] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, San Francisco, CA, 2001.

[24] P.E. Otto, G.B. Davies, N. Chater, and H. Stott, "From spending to understanding: Analyzing customers by their spending behavior," Journal of Retailing and Consumer Services, vol. 16, no. 1, 2009, pp. 10- 18.

[25] Hand, D. J., Whitrow, C., Adams, N., Juszczak, P., and Weston, D.,"Performance criteria for plastic card fraud detection tools" : Journal of the Operational Research Society, vol. 59, no. 7, 2008, pp. 956–962.

Leila Seyedhossein has received her B.Sc. degree in Electrical & Computer Engineering from the University of Shahid Beheshti in 1999. She is currently a M.Sc. student at the School of Electrical and Computer Engineering of the University of Tehran. Her research interests include data mining, fraud detection and e-commerce.

Mahmoud Reza Hashemi received his B.Sc. and M.Sc. in Electrical Engineering from the University of Tehran. He pursued his Ph.D. at the University of Ottawa, Canada. He is currently an assistant professor at the University of Tehran, and the Director of the Multimedia Processing Laboratory (MPL). His research interests include security of information technology services, fraud detection, peer-to-peer distributed e-marketplaces, multimedia processing and streaming. Dr. Hashemi has served on the program committees of a number of conferences in multimedia, e-commerce, and security.