A New User-Based Model for Credit Card Fraud Detection Based on Artificial Immune System

Neda Soltani
Faculty of Computer Engineering
and Information Technology
Amirkabir University of Technology
Tehran, Iran
neda.soltani@aut.ac.ir

Mohammad Kazem Akbari
Faculty of Computer Engineering
and Information Technology
Amirkabir University of Technology
Tehran, Iran
akbarif@aut.ac.ir

Mortaza Sargolzaei Javan
Faculty of Computer Engineering
and Information Technology
Amirkabir University of Technology
Tehran, Iran
msjavan@aut.ac.ir

Abstract—**In this paper we present a new model based on Artificial Immune System for credit card fraud detection. In this model, which is based on *Artificial Immune Recognition System*, user behavior is considered. The model puts together the two methodologies of fraud detection, namely tracking account behavior and general thresholding. The system generates normal memory cells using each user's transaction records, yet fraud memory cells are generated based on all fraudulent records. To get more accurate results, we have performed analysis on training data in order to control the number of memory cells. During the test phase each user's transaction is presented to his/her own normal memory cells, together with fraud memory cells.**

Keywords-Artifical Immune System; Credit card Fraud Detection; User profiling

I. INTRODUCTION

Credit cards are being used everywhere and have become a successful way of modern payment, while suffering from being misused. Using plastic cards in everyday payment activities, makes it easier for fraudsters to achieve novel ways of misusage. In this content we consider misusage as unauthorized account activity committed by means of the debit/credit facilities of a legitimate account [1]. Fraud detection is the act of recognizing such an activity and stopping it as soon as possible, namely before the transaction is accomplished. The related approaches are divided into two main subcategories. The absolute analysis that searches for thresholds between legal and fraudulent behavior, and the differential approach that tries to detect extreme changes in a user's behavior [2]. First approach is a supervised method in which we need fraud records to create the model and decide on thresholds. However, the second method is based on user behavior, which might use user profiling, behavioral models, and related methods. In this approach the transactions with a salient difference from normal behavior are flagged as fraud. Confirming whether a transaction was done by a client or a fraudster by phoning all card holders is cost prohibitive if we check them in all transactions. Fraud prevention by automatic fraud detections is where the well-known classification methods can be applied, where pattern recognition systems play a very important role. One can learn from past (fraud happened in the past) and classify new instances (transactions) [7].

According to [1] there are some challenges faced by a fraud detection system which stem from the nature of the transaction data and some particular operational issues:

- The number of transactions processed by plastic card issuers daily is high, furthermore each transaction includes more than 70 fields of coded information. Transaction data is heterogeneous and time-varying within and between accounts. Patterns and trends vary significantly for different groups of merchants, holiday seasons and geographical regions.
- The generally accepted fraud rate within the plastic card industry is 0.1–0.2%, i.e. the occurrence of fraud is relatively rare. Frequently this leads to the problem that the majority of cases flagged by the fraud detection system as being potentially fraudulent are in fact legitimate. This type of error is referred to as false positive (FP). As the number of FPs increase so do the associated costs and customer inconvenience.
- Alerts arising from the fraud detection system are usually passed on to the fraud department for further investigation. The suspected cases are followed up with a call to a cardholder for verification of the transactions, where it is required by the bank policy. As a result of this, the number of alerts should be kept at a level such that it can be handled by the available number of investigators and fraud analysts.
- Fraudulent cases missed by the fraud detection system are reported to the issuing company when the cardholder identifies that their account has been compromised. This can take up to several months, resulting in a delay in correctly labeling each case. Some fraudulent cases remain unidentified and therefore mislabeled. Thus, a fraud detection model is almost certainly trained on noisy data.

Fraud detection techniques which have been developed for a special field can be non-effective in other

Sponsor: Iran Telecommunication Research Center

978-1-4673-1479-4/12/$31.00 ©2012 IEEE

field, i.e. a typical credit card fraud detection system might not work well for detecting fraud in insurance systems and vice versa. So it is important to consider the features of the field. As mentioned before, there are two major approaches in credit card fraud detection. First, it is supervised classification which works in transaction level. In this method transactions are labeled as fraudulent or normal based on previous transactions. An example for such a system is rule based systems. The rules are extracted either by the fraud team, or by a tree-based algorithm. This approach is proven to reliably detect most frauds which have been observed before. The disadvantage of this approach is that there must be adequate number of a special fraud pattern in order to extract the rule. It costs time to collect all fraud patterns, extract rules, and put them in action and by that time fraudsters may have changed their tactics. Policies based on global thresholds have limited capabilities due to their inability to learn and adapt to observed account behavior commonly resulting in large volumes of false alerts to be resolved by a business analyst [5].

The second approach contains unsupervised methodologies; which observe account behavior and label the transactions which are in contrast with user's normal behavior in a time window. This is because we don't expect fraudsters behave the same as the account owner or be aware of the behavior model of the owner [3]. One example of this approach is behavioral models which creates a user profile for each account. This profile contains of the activity information of the account; such as merchant types, time of transactions, amounts, locations, etc. It is impossible to cover all legal behaviors as there are lots of transactions with different behavior patterns, and the opportunity to accept new behavior patterns. This method has a high rate of misclassified normal transactions because it labels any transaction different from the user's past behavior while it might have been done by user him/herself.

In conclusion both approaches have drawbacks. The first approach, which is based on misuse detection, cannot detect new patterns of frauds, and the second one, which is based on anomaly detection, has many wrong alarms [3]. So one can think of using both approaches.

II. Artificial immune systems

A. *Artificial Immune System*

The natural immune system is a highly complex system, comprised of an intricate network of specialized tissues, organs, cells and chemical molecules. The natural immune system can recognize, destroy, and remember an almost unlimited numbers of pathogens (foreign objects that enter the body, including viruses, bacteria, multi-cellular parasites, and fungi). To assist in protecting the organism, the immune system has the capability to distinguish between self and non-self. Notably, the system does not require exhaustive training with negative (non-self) examples to make these distinctions, but can identify items as non-self which it has never before encountered [8].

Artificial Immune Systems are designed based on human body immune system which is made up of various cells having different functions. All organs are distributed in human body and there is no central point controlling them. All the elements are moving along body acting complementary with each other. The main function of immune system is to seek malfunctioning self cells (like cancer cells), and pathogenic non-self cells (like viruses and bacteria). Any detected cell by the immune system is named antigen. The cells which are self and do not hurt body are self antigens, and those which cause disease are non-self antigens. The immune system must be able to discriminate between self and non-self. This is done by lymphocytes, which is called detectors in this paper. Detectors are created based on random protein patterns. Then they go through a process named Negative Selection. In this process any detector which detects a self cell, is omitted (actually dies). The detectors which survive this operation are entered into blood and start questing non-self cells. So the detectors become tolerant to self. Then the detection phase starts. Detectors leave a couple of days and if they don't match any non-self antigen, they die. If any detector comes across a non-self antigen and has a high affinity with it, it makes some clones. This cloning process contains mutation. The created clones might have higher affinity with the non-self antigen, so the best matching clone is selected as memory cell and leaves longer. This process is named Clonal Selection and tries to keep the best detectors to detect the same patterns of non-self antigens later.

B. *Related Work*

There are some publications about using AIS for fraud detection, some of which address credit fraud detection. The results are promising but still there are some drawbacks for each method. In [9] the authors propose a method based on vaccination along with negative selection. Vaccination is very similar to clonal selection. The results are presented for different fraud types. While the method works well detecting some fraud types, it fails to detect some types with high precision. Yet the overall results are promising. In [8] the authors propose using Distance Value Metric for calculating distance between records. This metric is based on probability of the occurrence of values in the training set. The results show increase in detection rate. But still the number of false alarms, and missed frauds are high. Clonal selection is compared to negative selection, which shows that clonal selection is more accurate having higher detection rate. [10] addresses fraud detection in VoD systems. The proposed method combines two algorithms in the field of AIS with a regression tree. The results are promising. Though we cannot conclude that the method will work as a credit card fraud detection as well; because as mentioned before, a fraud detection system must be implemented considering the field and its characteristics. In [7] a comparison has been done between different methods for credit card fraud detection. The results show that after improving input parameters for all the methods AIRS has the best results. This is because the number of input parameters for AIRS is relatively high. If we consider training dataset, and set the parameters depending to dataset, the results tend to become better.

There are not many works done in this area which report good results. AIS has been used for intrusion detection so far. But fraud detection using AIS, though it is not a new idea, has not been taken into consideration very much. Still the characteristics of AIS, and the similarity between two fields implies that there can be good results.

C. Artificial Immune Recognition System

AIRS is a classification algorithm which is based on AIS. It uses clonal selection to create detectors. AIRS generates detectors for all of the classes in the dataset and in detection phase uses k Nearest Neighbor algorithm to classify each record. Affinity is calculated using Euclidean distance. So the k neighbors are the records which has the least Euclidean distance with the test record.

An artificial immune system (AIS) is a class of adaptive or learning computer algorithm inspired by function of the biological immune system, designed for and applied to difficult problems such as intrusion detection, data clustering, classification and search problems. It is critical at the outset to stress that although terminology and function of AIS are described using biological terms from the field of immunological research, they are taken as simplifications and abstractions and not intended to be models or representative of immunological response systems. The field of AIS research has been around for approximately 15 years, and for the majority of its history has been concerned with feature extraction (data clustering), and change or anomaly identification, such as intrusion detection systems. More recently, AIS has been applied to broader domains such as function optimization, and in the case of AIRS, classification. The recent shift in applicable problem domains has required a rethink of the algorithms application and adaptation of existing tried and tested AIS elements. This new field of research provides an opportunity for innovation, not only in the designing new specialized AIS algorithms, but also in the successful encoding the adaptive power of the metaphor.

The AIRS algorithm was one of the first AIS technique designed specifically and applied to classification problems. It has been shown to exhibit the following desirable algorithmic characteristics [4]:

- Self-regulation – A problem common to the field of artificial neural networks is the selection of an appropriate topology or neuronal architecture. AIRS does not require the user to select an architecture, instead the adaptive process discovers or learns an appropriate architecture during training.

- Performance: Results of AIRS compared to the empirical results of the best known classifiers show that AIRS is a competitive classification system. Results indicate that AIRS can achieve classification accuracy in the top five to top eight when ranked against some of the widely known best classification systems is capable of achieving the best classification result known for some datasets.

- Generalization: Unlike techniques such as k-Nearest Neighbor that use the entire training dataset for classification, AIRS performs generalization via data reduction. This means that the resulting classifier produced by the algorithm represents the training data with a reduced or minimum number of exemplars. It is typical for AIRS to produce classifiers with half the number of training instances.

- Parameter Stability: The algorithm has a number of parameters that allows tuning of the technique to a specific problem, with the intent of achieving improved results. A feature of the algorithm is that over a wide range of parameter values, the technique is capable of achieving results within a few classification accuracy percentage points of the results achieved with an optimal parameter set.

III. Methodology

As mentioned before, fraud detection methods are based on two approaches: user behavior analysis, and fraud analysis. AIRS can be classified in the second group. As it does not consider user behavior and is a supervised learning method. In this paper we propose adding the first approach to AIRS in order to get more precise results. Fraud detectors will demonstrate the fraud patterns in the dataset. While legal transaction detectors will be generated specific for each user, and demonstrate user behavior. This is because users have different shopping habits, and we cannot consider a unified pattern for all users or cover all the possible scenarios for all users. In addition, one transaction done by a user can be considered normal, while the same transaction done by another user is a fraud. So we have to differentiate between each user's detectors.

In the proposed approach we generate normal detectors for each user. This means while choosing the best memory cell as the normal detector for a transaction of a user, only the same user's previous transactions are considered and processed. Yet the fraud detectors are generated based on all fraudulent transactions in order to cover all possible fraud types. When testing a special user's new transaction to determine its class (either normal or fraud), the k neighbors are chosen from all fraud detectors, and the same user's normal detectors.

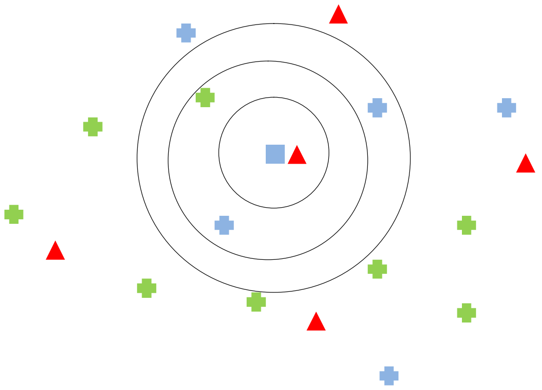

Figure 1. How user-specific detectors work while testing a record

Figure 1 shows how the proposed method works. Plus shaped figures demonstrate normal cells, and triangles demonstrate fraud cells, and the rectangle demonstrates the test record. We have separated users of cells and records by using different colors for each user. So all blue shapes belong to one user, and all green shapes belong to another user. If value of k in the KNN algorithm is equal to 1, the first circle is the boundary in which the cells are considered for classifying the record. If k is 3, unmodified AIRS will consider 2nd circle. But user-specific AIRS won't accept the green plus because it is a cell for a different user. So the boundary is the 3rd circle. This is how the classification process takes place.

A. Analysing Dataset

There is another issue about getting the best results on a specific dataset. Records in different datasets have different fields and specifications. The volume of fraud in every dataset is different. This might be because of the different security protocols used by different organizations and banks. Whatever the reason is, this fact causes different fraud characteristics on each dataset, which affects the performance of the fraud detection system. Therefore considering the dataset characteristics before generating detectors will help the system in generating the best detectors and having more precise results. For this reason we have added a dataset analysis function to the training phase. In this function all fields are examined, and *alpha* index is calculated for each field using (1):

$$alpha = N_{ij} / N_f \qquad (1)$$

$alpha_{ij}$ is the fraud ratio for the *j*th value of *i*th field, and N_{ij} shows the number of fraudulent records having *j*th value for the *i*th field, and N_f is the whole number of fraudulent record in the dataset.

IV. Results and discussion

We have tested our proposed method on a dataset of transactional records, which have been registered by a POS. The transactions have 12 fields. Sensitive fields have been anonymized. We ignored some fields like time details which contained the time of transaction in seconds. Remaining fields contain customer code, date and time of transaction, location, status, merchant code, and amount. The dataset is labeled (fraudulent records are labeled as fraud) and contains 0.2% fraudulent transactions. We trained the method on 70% of the dataset and the remaining 30% was used for test. We used WEKA for testing the method. We used *wekaclassalgos* plugin which contains immune algorithm including AIRS. The changes have been done on the java code in this plugin.

A. Results

The results of the tests are shown in Table I. The evaluation metrics are as follows:

Detection Rate = TP / (TP+FN)

FP Rate = FP / (FP+TN)

FP Ratio = FP / TP

First line in Table I shows unmodified AIRS algorithm. Second line shows AIRS in which we have created memory cells for each user separately. The third line adds the *alpha* index to the training phase. The first column shows detection rate. While AIRS cannot detect any fraud, using user specific memory cells makes a big difference. Detection rate increases up to 91%. This is a high detection rate, and also a big increase. But the method increases the FP rate too. As shown in Table I, FP has a high rate which shows from every 3 normal transactions one is labeled as fraud. Though we see a high increase in detection rate the precision decreases. In the third column we can see the FP ratio which shows the ratio of FP over TP. This column shows that the number of FPs is more than the number of TPs. That's why we cannot claim this method works better than AIRS. The third line of the table has more promising results. While the detection rate increases up to 100%, FP rate decreases by 15%. Though the FP rate is more than unmodified AIRS, this increase can be ignored as the detection rate is maximum. Considering the fact that FN imposes much more cost that FP does, we can conclude that this increase in FP is ignorable.

TABLE I. Results of AIRS and proposed methods on the whole dataset

Method	**Metrics**		
	Detection Rate%	***FP Rate%***	***FP Ratio***
AIRS1	0	0	0
AIRS1-User Specific Cells	91	33	214.7
AIRS1-User Specific Cells; Dataset considered	100	13	98.9

$\alpha + \beta = \gamma$ (1) (1)

In these test we considered the whole dataset containing all users. Some users had only less than 5 records in the dataset. This could be a reason for false alarms. Extracting a behavior through fewer numbers of transactions does not make sense. So in a second test we selected transactions of only those of users who have more than 10 records in the dataset. So the cells which are generated for each user is much closer to his/her overall behavior. The results of the second test are shown in Table II.

TABLE II. Results of AIRS and proposed methods on the selected users of dataset

Method	**Metrics**		
	Detection Rate%	***FP Rate%***	***FP Ratio***
AIRS1	33	0.09	1.67
AIRS1-User Specific Cells	100	10.30	66.67
AIRS1-User Specific Cells, Dataset considered	100	9.89	64

The second test shows the same trend between several methods. User-specific AIRS gains more detection rate and less FP rate than the first test. Yet it is much better than unmodified AIRS while considering detection rate. The proposed method, which contains user-specific cells and *alpha* index, has less FP rate than the second method. The difference between 2nd and 3rd lines of the table has become less than Table 1. The overall differences with the first test are because of having a more realistic dataset in this test.

B. Conclusion and Future work

Considering user behavior results in detection of more fraudulent transactions, while at the same time has a higher false alarm rate. This is because of the different behavior of users in different time windows. Sometimes users do not tend to have a specific type of behavior; which is why a user-based method in fraud detection has more false alarms. A solution to this issue is to update user model based on his/her latest transactions. The fact that the dataset did not contain plenty of transaction records for each user is another reason for this. There were users with only one or two transactions which obviously cannot present user's behavior.

One desirable characteristic of AIS is adaptable learning. Those detector cells that fail to interact with an antigen die after a few days. But cells that do interact are stimulated, triggering an immune response. They immediately begin dividing, producing more cells that can detect the same antigen [6]. One way to decrease the number of false alarms is to delete the detectors which have not been used to detect any cells for a long time. This is done by counting each time a detector rests between K neighbors of any record. Then the detectors which count zero times will be ignored or deleted.

REFERENCES

[1] M. Krivko, "A Hybrid Model For Plastic Card Fraud Detection Systems", Expert Systems with Applications, Vol. 37, No. 8, pp. 6070-6076, 2010.

[2] S. Hilas, "An application of supervised and unsupervised learning approaches to telecommunications fraud detection", Knowledge-Based Systems, vol. 21, No. 7, pp. 721-726, 2008.

[3] A. Kundu, S. Panigrahi, S Sural, and A. K. Majumdar, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection", IEEE Transactions on Dependable and Secure Computing, Vol. 6, No. 4, October-December 2009.

[4] J. Brownlee, "Artificial Immune Recognition System (Airs) A Review and Analysis", Technical Report, Victoria, Australia: Centre for Intelligent Systems and Complex Processes (CISCP), Faculty of Information and Communication Technologies (ICT), Swinburne University of Technology, January 2005.

[5] M. Edge, and P. Falcone Sampaio, "A survey of signature based methods for financial fraud detection", Computers & Security, Vol. 28, No. 6, pp. 381-394, 2009.

[6] E. Klarreich, "Inspired by immunity." Nature Vol. 415, pp. 468-470, 2002.

[7] M. Gadi, X. Wang, A. Lago, "Credit Card Fraud Detection with Artificial Immune System", Springer, 2008

[8] A. Brabazon, et. al., "Identifying Online Credit Card Fraud using Artificial Immune Systems", IEEE Congress on Evolutionary Computation (CEC), Spain, 2011

[9] N. Wong, et al., "Artificial immune systems for the detection of credit card fraud: an architecture, prototype and preliminary results", Information Systems Journal, Vol. 22, No. 1, pp. 53–76, 2012

[10] A. Watkins, J. Timmis, and L. Boggess, "Artificial immune recognition system (AIRS): An immune-inspired supervised machine learning algorithm", Genetic Programming and Evolvable Machines, Vol. 5, No. 3, pp. 291-317, September 2004

[11] Andrew Watkins, Jon Timmis, "Artificial Immune Recognition System (AIRS): Revisions and Refinements," 1st International Conference onArtificial Immune Systems (ICARIS2002), University of Kent atCanterbury, pp. 173-181, 2002.

[12] A. Watkins, J. Timmis, and L. Boggess, "Artificial immune recognition system (AIRS): An immune-inspired supervised machine learning algorithm", Genetic Programming and Evolvable Machines, Vol. 5, No. 3, pp. 291-317, September 2004

[13] W. Albrecht, C. Albrecht, and et al, "Current Trends in Fraud and its Detection", Information Security Journal: A Global Perspective, Vol. 17, No. 1, pp. 2-12, 2008.

[14] A. Kundu, S. Panigrahi, and et al, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection", IEEE Transactions on Dependable and Secure Computing, Vol. 6, No. 4, pp. , 2009.

[15] A. Srivastava, A. Kundu, and et al, "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on dependable and secure computing, Vol. 5, No. 1, pp. 37-48, 2008.

[16] D. Sanchez, M. Vila, and et al, "Association rules applied to credit card fraud detection", Expert Systems with Applications, Vol. 36, No.2, pp. 3630-3640, 2009.