Journal of Advances in Computer Research
Quarterly pISSN: 2345-606x eISSN: 2345-6078
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 8, No. 2, May 2017), Pages: 67-77
www.jacr.iausari.ac.ir

Providing a Method to Identify Malicious Users in Electronic Banking System Using Fuzzy Clustering Techniques

Leila Pourabdi[✉1] **and Ali Harounabadi**[2]
1) Department of Computer Engineering, Saveh Branch, Islamic Azad University, Saveh, Iran
2) Department of Computer Engineering, Tehran Markaz Branch, Islamic Azad University, Tehran, Iran
l.poorabdi@gmail.com; a.harounabadi@yahoo.com

Received: 2016/06/23; Accepted: 2016/08/16

Abstract

Money-Laundering causes a higher prevalence of crime and reduces the desire tending to invest in productive activities. Also, it leads to weaken the integrity of financial markets and decrease government control over economic policy. Banks are able to prevent theft, fraud, money laundering conducted by customers through identification of their clients' behavioral characteristics. This leads to reduce the banking and credit risks. So there are some systems in order to identify unusual users' behavior in banking industry that can help different societies. In present study, effective variables are used to determine suspicious behavior in terms of money-laundering from users' account transactions in an Iranian private bank. Users' membership degree to clusters is determined using fuzzy clustering method and maximum membership degree is considered as a label for users; also, back propagation neural network is used to identify the model. The results show that the proposed method can detect money-laundering accurately at the bank up to 97%.

Keywords: *Money-Laundering, Fuzzy Clustering, Membership Degree, Neural Network*

1. Introduction

In general, "money laundering" or "laundering of contaminated money" is a process that the origin and characteristics of contaminated money from illegal trades such as drug trafficking, organized crime, terrorism and such these are changed and taken in an illegal form. Due to the dramatic increase in crimes and criminal acts, money laundering has been growth globally. It is obvious that these activities will cause to irreversible losses to the country's economy. So that according to estimation of the International Monetary Fund and the World Bank, illegal revenue being in the cycle of money laundering is approximately %2to %5 of global gross product and such damage would require the government to deal with this phenomenon[1]. Money laundering is a three-stage process; for laundering the revenue of crime, three distinct stages are identified: The Deployment Stage, The Layering Stage, and The Integration Stage.

1.1 The Placement Stage

The first stage of money laundering is placement. This stage is the most dangerous for brokers because probability of crime detection is very high at this stage. Depositing a substantial amount of cash to one or more accounts in banks will be taken into many people's consideration. One of the most common methods applied by money launderers is employment of people who deposit the cash into bank accounts for several times and less than specified limit for reporting, and they mainly use the small or a combination of different notes. Also, it is possible that these people receive traveler's check or payment order. The tool will then be deposited in the target bank. Substitution through electronic draft, conversion of currency or stock or securities (etc), property purchase, speculation and mass displacement, gambling and betting, payment order, insurance purchase and etc are the other methods of money laundering at this stage.

1.2 The Layering Stage

The second stage is layering or overlying. In this stage, contaminated money of financial institution established at the beginning was withdrawn and displaced through some other financial institutions and applied in complex transactions. Using high-value credit cards to pay foreign purchase and deposit funds to unknown banks in "tax haven" are methods applied by money launderers and tax evaders. In general, the purpose of money launderers at this stage is that the identification of contaminated money gets more difficult by the government and tax officers. Electronic transfer of money, transferring money to shopping institution, deposits, current accounts and other intermediary are the other ways in which they are used by money launderers.

1.3 The Integration Stage

The last stage of money laundering is integration. This occurs when laundered money can be restored to use the legal system that it may not be identified. Re-loan, loan default, collateral loan etc. are some methods used in this stage. One of the other efficient methods in this stage is acquisition buying of a bank in tax havens. Ownership of these institutions allows money launderer to change all signs of contaminated money by doing various activities. Credit card, buying financial assets and resell them, export and import, establishing financial institutions, etc. are some ways that money launderer apply to be able to restore contaminated money towards legal system [2].

2. Related Works

Banks by knowing behavioral characteristics of their clients can prevent theft, fraud, money laundering conducted by customers. This leads to reduce the banking and credit risks. Nonparametric methods and data mining like decision trees, neural networks and expert systems were applied. In this study, some proposed recommendation is studied to identify fraud. In the following, there are a few other methods for the detection of crimes is considered in the banking system. Fraud is a serious crime that it necessitates the development of transactions detection methods. Some studies were conducted, but the problem is not solved completely. ARadial Basis Functions(RBF) neural network model based on clustering algorithm APC-III and subtractive algorithm of least squares to conflict money laundering was proposed. Clustering algorithm APC-III is used to determine the parameters of RBF in hidden layer and Regression Least Squares (RLS) algorithm is adopted to update links weights between the hidden layer and output layer.

Since the RBF neural network can perform calculations from time to time by their units, they can determine whether the flow of capital is included in money laundering activities [3]. A clustering algorithm combining k-means and Birch and decision tree is used to detect money laundering. The main purpose of decision tree is classification of data objects through a top-down approach. Many clustering algorithms are ineffective in very large databases and it is not considered the appropriate memory for cases which are too large in database. Birch is an unsupervised data-analysis algorithm to perform hierarchical clustering in large data set in particular, and an incremental algorithm. But it cannot control financial data well. But Birch is sensitive to noise and derived data which is the most important part in anti-money laundering. In addition, K-means algorithm can easily handle financial data, but it is so costly. Liu and his colleagues suggested the following decision tree for their system based on Birch and k-means to solve the problems [4]. Id3 algorithm is used in order to assess the risk of money laundering based on decision tree. The most important feature is selected as the root of tree and the rules are derived from the root to some leaf nodes. Each path from the root to the leaves provides a classification rules from examples. Each path from root to leaves provides a rules' classification of cases. Using their features and ranges through the extracted rules of the branches, some features were identified as the most important in money laundering [5].

A model of graphs that may include suspicious transactions has been created. The model used in this paper is parameterized. Using fuzzy numbers indicates the parameters of transactions and parameters of discovered transaction graphs. Since money laundering may include the transfer of money through various accounts, the presented graphs transactions model parameterized due to some structural features [6].

A association rules are used which are based on Learning the specific rules and are able to discover indicators of deceptive behaviors from large databases of user's transactions. These indicators are applied to provide monitor system in order to record unusual behavior of customers and identify suspicious ones. Finally, the output of these systems can be used to alarm and warn the offending users [7].In order to detect crime in cash-card, it is used fuzzy logic and Takagi-Sugeno Neural-Fuzzy to Deutsche Bank Fraud Transactions. Using rules generated by fuzzy logic and based on professional experts' idea, it is able to detect suspicious users' behavior in terms of their fraud[8].

3. Fuzzy Clustering

The data may contain complex structures that even the best data mining techniques are not able to extract significant patterns from them. Clustering provides a way to identify complex data structures and separate inconsistent competitive signals to their parts. Clustering is called to division of heterogeneous population into a number of subsets or homogeneous clusters. The basic idea in fuzzy clustering is such that we assume each cluster is a set of elements, then by changing in definition of elements membership in this set it can be a mode of an element which is only able to be from a cluster member (partition mode) to a mode which each element can be set with different degree in several clusters, the classification that are more compatible with reality. Fuzzy C-Means (FCM) is one of the most popular techniques of clustering algorithm [8]. FCM clustering algorithm depends on many factors such as the number of clusters and determination of the distance between the clusters. One of the most important issues in clustering is selection of a number of appropriate clusters. It can be said that the number of clusters is useful that [10]:

Samples in a cluster are similar to each other as much as possible; common criteria for determining the value are data density and variance data. Samples of different clusters are separated as much as possible.

For FCM algorithm, during the data clustering, the value of all features is considered equal. While the actual data collection, some features may be more significant than other. For this, during the clustering some feature having more power and value should be significantly considered more. To solve this, some methods have been proposed in which each feature is evaluated based on its weight. In recent years, various clustering algorithm is provided with features weighting. In this paper, the proposed algorithm Improved Feature-Weighted Fuzzy C-Means(IFWFCM) is used [11].

Imagine that we have N training data that each this data has D different dimensions.

Also we have the fuzzy membership matrix indicating each data belonging to each cluster and c represent each cluster. In order to minimize the objective function, this algorithm is as follows:

$$J_m(U,V,w:D)=\sum_{i=1}^{c}\sum_{j=1}^{N}\left(\mu_{ij}\right)^m\left[d_{ij}^{(w)}\right]^2 \tag{1}$$

Here $d_{ij}^{(w)}$ is weighted Euclidean distance using the weights $w=\left(w_1,w_2,...,w_d\right)^T$ for features that are defined as follows:

$$d_{ij}^{(w)}=\left\|diag(w)\left(x_j-v_i\right)\right\| \tag{2}$$

Here diag(w) is defined as follows:

$$diag(w)=\begin{pmatrix} w_1 & 0 & ... & 0 \\ 0 & w_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & w_d \end{pmatrix} \tag{3}$$

It should also be noted that the following terms must be met in these equation.

$$\sum_{i}^{c}\mu_{ij}=1 \tag{4}$$

$$\sum_{q=i}^{d}w_w=1 \tag{5}$$

$$J_m(U,V,w:D)=\sum_{i=1}^{c}\sum_{j=1}^{N}\left(\mu_{ij}\right)^m\left[d_{ij}^{(w)}\right]^2-\lambda\left(\sum_{i=1}^{c}\mu_{ij}-1\right)-\beta\left(\sum_{q=1}^{d}w_q-1\right) \tag{6}$$

Where λ and β are Lagrange multipliers; finally, using Lagrange function, clustering centers updating formulas, membership functions as well as features weight vector is calculated as follow:

$$\mu_{ij}=\frac{1}{\sum_{k=1}^{c}\left(\frac{\left[d_{ij}^{w}\right]^2}{\left[d_{kj}^{w}\right]^2}\right)^{\frac{1}{(m-1)}}} \tag{7}$$

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^{m} x_j}{\sum_{j=1}^{n} \mu_{ij}^{m}} \tag{8}$$

$$w_q = \frac{1}{\sum_{L=1}^{d} \left[\frac{\sum_{i=1}^{C} \sum_{j=1}^{N} \mu_{ij}^{m} (x_{jq} - v_{iq})^2}{\sum_{i=1}^{C} \sum_{j=1}^{N} \mu_{ij}^{m} (x_{jL} - v_{iL})^2} \right]} \tag{9}$$

Clustering pseudo-code IFWFCM for clustering data is as follows:

Input: Dataset $D = \{X_j\}_{j=1}^{N}$
Output: Terminal fuzzy partition matrix $U(t)$
Terminal center matrix $V(t)$
Terminal feature- weight vector w
Initialization: Feature-weight vector w,
Number of clusters C,
Fuzzification exponent m,
Fuzzy partition matrix U
For i=1to $\|U^{i+1} - U^{i}\| \le \varepsilon$
Step1:Compute the cluster center
Step2:Calculate the weighted distances
Step3:Update the fuzzy partition matrix
Step4:Update the elements in the featureweightvector
End For

4. Validation and Determination of Optimal Number of Clusters

Fuzzy clustering algorithm receives the number of clusters as input. So before the start of the algorithm, it is necessary to know the optimal number of clusters in a dataset. For this, clustering validity indices are used.

Partition Coefficient:

The division multiplier index which is a simple one is defined as follows:

$$V_{PC}(U) = \frac{1}{n} \left(\sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^{2} \right) \tag{10}$$

$$Max \left\{ \left(V_{PC}(U, c_i, m) \right) \right\} \tag{11}$$

Where "C" and "N" refer to number of cluster and data number respectively; u_{ij}^2 indicates degree of data membership. $Max \left\{ \left(V_{PC}(U, c_i, m) \right) \right\}$ indicates to the most optimal number of clusters when V_{PC} is at its highest value[12].

The amount of V_{PC} than m has uniform dependence. Diverse levels are recommended in (1,30] for m, but m<5 is normal and m=2 is more common [13].

Pal and Bezdek in [13] until find the most optimized cluster level tested V_{PC} for different m, in 2≤c≤ 10. The number of cluster which has the highest V_{PC} level was selected as the most suitable cluster number.

5. Neural Network Structure

Neural network includes components of layers and weights and network behavior depends on relationship between members as well. In general, there are three types of neuron layer in neural networks:

Input layer: This layer receives inputs and it sends the input signal to the next layer in terms of its relationship power.

Middle layer (hidden layer): The number of middle layers and their neurons is optional; this layer should be selected carefully to give appropriate output. The function of these layers is determined by inputs and link weight between them and hidden layers. Weights between input and hidden units reveal that when a hidden unit should be activated.

Output layer: another group of neurons makes the outside world through its output; the performance of output unit depends on hidden unit activity and the link weight between the hidden and output units.

Learning a Back-Propagation (BP) network consists of three phases: pre-entry input learning pattern, related error Back-propagation and adjusting weights. Propagation learning algorithm is that for a given input pattern, it provides an output network and compares this reaction and optimal reaction of each neuron. Network weights are then modified to correct or reduce the errors and the later pattern is shown. Weighing modified continuously continues in this process until all errors of predetermined tolerance level are minimized. The purpose of learning is that the weights are adjusted in such a way that was achieved through providing sets of optimal input-outputs. To do this, the network is usually learned using a large number of input-output called “cases”.

6. Dataset

Applied data collection has 5 properties related to users in a 5-month transaction that we are able to identify kind of users’ behavior in terms of suspecting to money laundering factor in a specific period of time. To do this, 999 users were selected. Applied properties in this study include: The standard deviation of the client's account in a period of time, having a joint account, the number of reciprocating transactions between several accounts, Number of functions more than the threshold, the number of additional accounts at the same bank. These features for all withdrawals and deposits are as physical presence, a Check, Internet Bank, Atm, etc. There was 3 Class in data set: The first cluster users with normal behavior, the second cluster users with little suspicious behavior and the third cluster users with suspicious behavior.

7. Method

In this study, we tried to be able to cluster users due to their traits through users’ account data that it is possible by settlement of a large number of money into one or more accounts according to many transactions and more functions than normal level that at the first level of money laundering which is the most important phase for identifying; and also fuzzy clustering which has more accurate results than classic clustering. Neural Network was used to identify new users' trait, too; after adequate training, the plexus could allocated users to that classes which are obtained clusters at clustering phase.

Partition coefficient is used to determine the optimal number of applied clusters. Partition coefficient is a validation criterion that its highest value indicates to optimal number of clustering.

In this study, V_{PC} index for limited number of m was obtained. As we can observe in table 1 and figures 1, 2, 3 and 4, V_{PC} index represented c=3 for different m in clusters in $2 \leq c \leq 10$.

Table 1. V_{PC} ***index for different m***

Number of cluster (c)	m=1	m=1/2	m=2	m=13
2	1	0/93	0/9	0/34
3	**1/04**	**0/98**	**0/95**	**0/38**
4	1	0/93	0/9	0/34
5	0/96	0/9	0/86	0/3
6	0/91	0/86	0/82	0/26
7	0/78	0/73	0/68	0/11
8	0/69	0/63	0/59	0/02
9	0/88	0/82	0/75	0/08
10	0/65	0/62	0/56	0/02

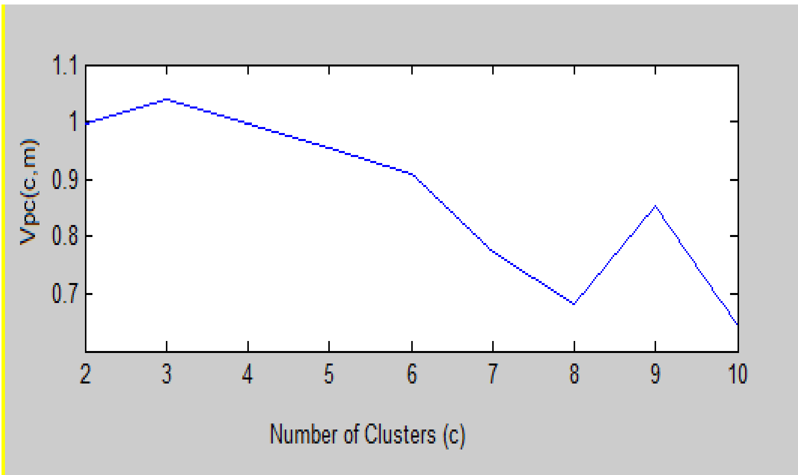

Figure 1. V_{PC} ***index level for m=1***

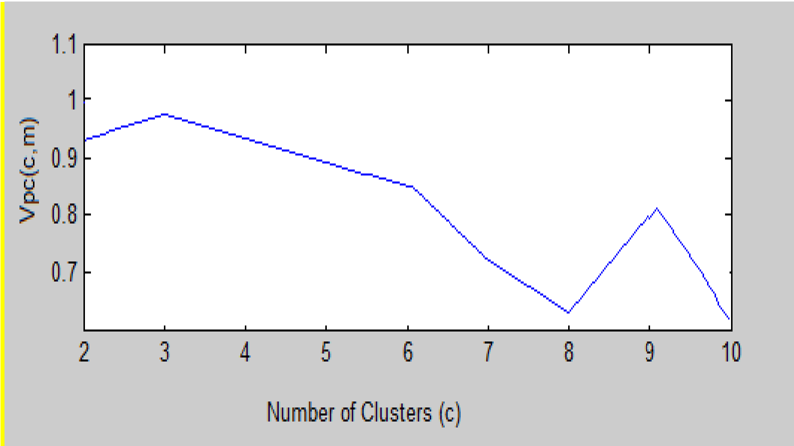

Figure 2. V_{PC} ***index level for m=1/2***

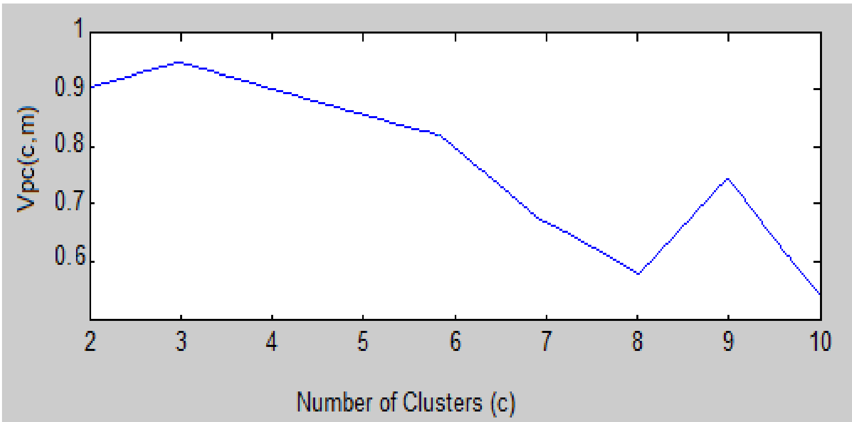

Figure 3. V_{PC} ***index level for m=2***

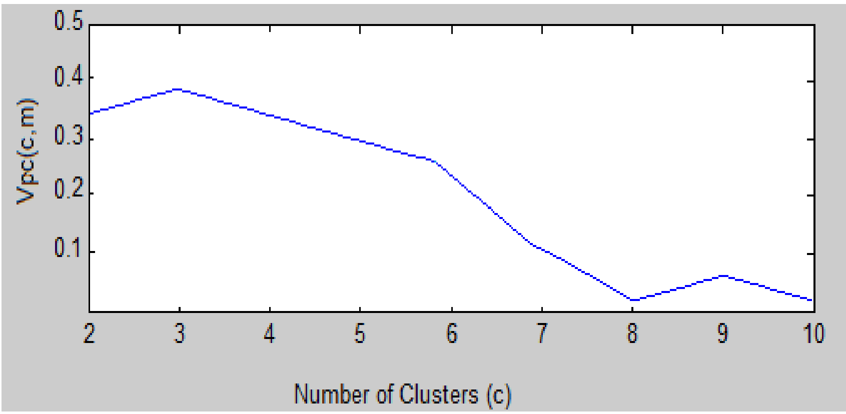

Figure 4. V_{PC} ***index level for m=13***

After determining the number of clusters, these are taken to clustering algorithm IFWFCM. That is, it is determined C=3 in clustering algorithm. After fuzzy clustering IFWFCM, it can be achieved the membership degree of each user in different clusters using membershipfunction matrix.

	1	2	3	4	5	6	7
1	0.9125	0.9845	0.0960	0.0960	0.9845	0.0960	0.6646
2	0.0857	0.0151	0.9023	0.9023	0.0151	0.9023	0.3328
3	0.0019	3.4469e-04	0.0017	0.0017	3.4469e-04	0.0017	0.0026

Figure5. The membership of users in each of the three clusters

According to Figure 5, membership degree of each user was obtained in 3 clusters. Therefore C=3 indicates to the number of output layer in neural network. The first user with membership degree of 0.9125, 0.0857, and 0.0019 are classified as first, second, and third clusters respectively. So the first user with the membership rank 0.9125 belonged to cluster users with normal behavior. Clusters of users were obtained and the maximum degree of membership to the cluster was considered as a label for each user. K-fold cross validation assessment was used to validate the data. Due to data K=5 and k-1 parts and one part were used to train and assess or test model, respectively.

80%users are selected for training and 20% for test data. Since we have 5 properties, the input layer neuron is 5. Also, there are hidden neurons in middle layer. In this study, 20 neurons are used in middle layer. Training rate is considered 0.001, so 280 frequencies are obtained. Stimulation function of tangent middle and output layers are hyperbolic and sigmoid respectively. Using training data having been labeled, we could train Back-propagation neural network and the accuracy of neural network is evaluated by 20% unlabeled test data. The accuracy of neural network is obtained97 % in order to allocate members.

8. Conclusion

The purpose of this study is to present a method to identify suspicious behavior of users in electronic banking. Important behavioral characteristics are selected from bank data collection and classified by clustering algorithm IFWFCM and users division index PC with same behaviors in same clusters. Finally, Back-Propagation neural network was used due to having learning ability to identify new users' behavior. Experiments show that the proposed method is more accurate in identifying suspicious users' behavior. Study on previous works is shown that whereas there is less possibility to reach labeled data, the process of researches trends to unsupervised method. Another achievement of this study is that it can be determined the number of clusters using validation methods and then, clustering algorithm is used for clustering.

9. Performance Assessment

We have simulated proposed model performance results with the results obtained from FCM clustering method and BP neural network and compared. Totally, the results have been obtained on the used data sets in this article. Stimulation has been done using MATLB R2013b. The results of conducted methods are shown in Figure 6:

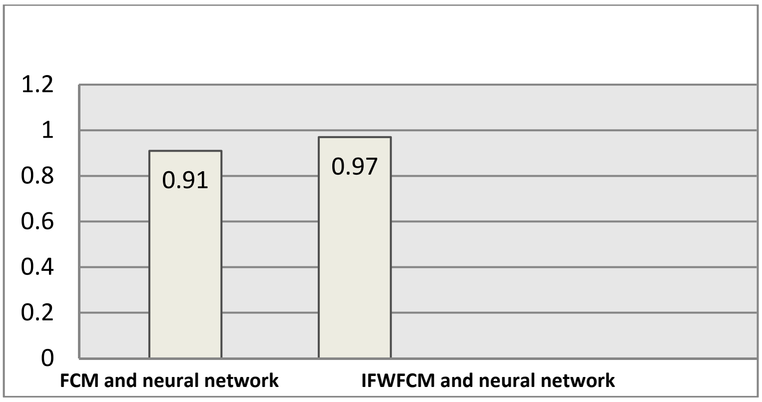

Figure 6. accuracy of proposed methods and FCM with neural network

10. Future Works

From present study, it is recommended the following suggestion for future research: it can be used the other features of money laundering to identify suspicious behaviors such as this important feature which the high number of deposits with dim sum is a considerable amount in a short time. When data label was not available for security reasons, clustering algorithms were used to identify data label. Using clustering and evolutionary algorithms are applied to improve accuracy. Improvement in existing clustering methods is done in order to enhance the accuracy of the clustering; that is, in weighted clustering when algorithm is used for the first time, weights are randomly determined for data; so it is better to use some methods such as Laplace score or data dispersion of each weight should be identified; since features' weight are different and the more important the features the higher weight it has.

References

[1] J.M.Myers,"International Strategies to Combat Money Laundering," Financial Crimes Enforcement Network. U.S. Department of the Treasury for the International Symposium on the Prevention and Control of Financial Fraud Beijing, 1998.

[2] R.J. Mechalowski, "Order, Law, and Crime: An Introduction to Criminology," New York : McGraw Hill, 1985.

[3] L.T. LV, N. Ji, and J.L. Hong, "A RBF Neural Network Model for Anti-Money Laundering," International Conference on Wavelet Analysis and Pattern Recognition, 2008.

[4] L. Rui, Q. Xiao-long, and M. Shu, and Z. Shuai-zheng,"Research on Anti-Money Laundering Based on Core Decision Tree Algorithm,"Control and Decision Conference (CCDC), 2011.

[5] W. Sunan, U. Zhejiang,andY. JianGang, "A Money Laundering Risk Evaluation Method Based on Decision Tree," Conference on Machine Learning and Cybernetics, 2007.

[6] K. Michalak, and J. Korczak,"Graph Mining Approach to Suspicious Transaction Detection," Conference on Computer Science and Information Systems, 2011.

[7] L. Fang, M. Cai, and H. Fu, and J. Dong,"Ontology -Based Fraud Detection," Conference on Computational Science, 2007.

[8] F. Shafiee Nezhad, and H. Shahriari, "Fuzzy Logic and Takagi-Sugeno Neural-Fuzzy to Deutsche Bank Fraud Transactions," 7 th International Conference on e-Commerce in Developing Countries with Focus on e-Security, 2013.

[9] A. Jain, M. Murty, and P. Flynn,"Data Clustering: A Review,"ACM Computing Surveys, 1999.

[10] T. Cura,"A particle swarm optimization approach to clustering," Expert Systems with Applications, 2012.

[11] X. Hong-Jie, and H. Ming-Hu, "Further improvements in Feature-Weighted Fuzzy C-Means,"Information Sciences, 2014.

[12] M. Ramze Rezaee, and B.P.F. Lelieveldt, and J.H.C. Reiber, "A new cluster validity index for the fuzzy c-mean," Pattern Recognition Letters, 1998.

[13] N.R. Pal, and J.C. Bezdek, "On Cluster Validity for the Fuzzy c-Means Model,"IEEE Transactions on Fuzzy Systems, Vol (3)3, 1995.