Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
June 2016, 23(3): 29–36
www.sciencedirect.com/science/journal/10058885
The Journal of China Universities of Posts and Telecommunications
http://jcupt.bupt.edu.cn

User abnormal behavior analysis based on neural network clustering

Zheng Ruijuan[1], Chen Jing[1] (✉), Zhang Mingchuan[1], Zhu Junlong[2], Wu Qingtao[1]

1. College of Information Engineering, Henan University of Science and Technology, Luoyang 471000, China
2. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract

It is the premise of accessing and controlling cloud environment to establish the mutual trust relationship between users and clouds. How to identify the credible degree of the user identity and behavior becomes the core problem? This paper proposes a user abnormal behavior analysis method based on neural network clustering to resolve the problems of over-fitting and flooding the feature information, which exists in the process of traditional clustering analysis and calculating similarity. Firstly, singular value decomposition (SVD) is applied to reduce dimension and de-noise for massive data, where map-reduce parallel processing is used to accelerate the computation speed, and neural network model is used for softening points. Secondly, information entropy is added to hidden layer of neural network model to calculate the weight of each attribute. Finally, weight factor is used to calculate the similarity to make the cluster more accuracy. For the problem of analyzing the mobile cloud user behaviors, the experimental results show that the scheme has higher detection speed (DS) and clustering accuracy than traditional schemes. The proposed method is more suitable for the mobile cloud environment.

Keywords anomaly analysis, information security, singular value decomposition (SVD), neural network, information entropy

1 Introduction

With the rapid development of mobile cloud computing, the mobile cloud service [1] business is bound to explosive growth. Therefore, people begin to store all kinds of information from computers to the cloud to reduce the constraints of its limited resources such as storage and computing resources. It brings convenience to work and life for people while also makes information security face severe tests.

With all kinds of hackers and intrusion behaviors emerging in endlessly, network attack technology becomes more mature and changeable, and the traditional passive defense means cannot solve mobile cloud user information security issues significantly. Facing of various passive defensive measures, people are more inclined to proactive detection techniques such as analysis to the abnormal behavior [2].

Received date: 12-11-2015
Corresponding author: Chen Jing, E-mail: 15036775207@163.com
DOI: 10.1016/S1005-8885(16)60029-8

Anomaly analysis [3] was proposed by James Anderson, whose main idea is modeling by some statistics of user behaviors to discover the ‘invaders’. The premise of abnormal analysis is assuming there are big differences between normal and abnormal behavior. The normal data is used to build model which can process the identifying data. If the matching results exceed the setting threshold, it will be regarded as an abnormal behavior.

Analyzing user abnormal behavior is actually the clustering problem [4], where the behaviors are clustered in two classes, i.e. “normal” and “abnormal”. Behaviors in the same class or cluster have higher similarity, while those in different clusters have lower similarity. The fact of abnormal analysis is how to divide the behaviors into several classes or clusters.

This paper merges together the ideas of SVD, neural networks and information entropy, which avoids the traditional clustering analysis problems of sensitive to noise and over-fitting. It uses information entropy and weight of attribute to calculate weight factor and similarity respectively, which makes the cluster more accuracy.

Based on the inherent defects of mobile terminals, this paper focuses on the abnormal behavior analysis method from user trusting aspect. The user requests will be received by wisdom mapping layer for further processing only when the user behavior is normal. Both analysis and simulation results indicate that our scheme outperforms other similar schemes.

2 Related works

Calculating similarity is very important to analyze user abnormal behavior. Several references in the field of user abnormal analysis have been summarized. The most common method is calculating the distance between sample properties. Ref. [5] used Euclidean distance to measure the similarity of attributes, which not only could measure the two-dimensional linear space, but also *d*-dimensional linear space. Ref. [6] used the lexical similarity *k*-means algorithm based on fuzzy logic Euclidean distance. It could improve the accuracy of the similarity estimation. Ref. [7] improved the *k*-means clustering algorithm performance, which uses the window technology in the process of clustering. Ref. [8] introduced segmental *k*-means algorithm into hidden semi-Markov model (HSMM) to train algorithm, where used the average information entropy of fixed-length observed sequence as the anomaly detection metric. These methods have two flaws. First, it cannot detect flood attacks. Second, it supposes that the weight of each attribute is equal, which floods the real weight of each attribute and greatly reduces the clustering accuracy.

The traditional clustering analysis is a hardening of the points, the classification category boundaries are distinct, which is likely to cause over-fitting. In fact, most of the objects have no strict attribute boundaries and suitable for softening points. Ref. [9] proposed the Bias-correction fuzzy clustering algorithms to avoid the hard clustering. It overcomes the poor clustering results and poor initializations. Refs. [10–11] calculated the information entropy of all records in the training dataset and the weight of each attribute with information entropy, which avoids the average and human interference to weight of each attribute and improves the clustering accuracy.

In order to make the analysis results more accuracy, Refs. [12–13] introduced neural network into the process of clustering, using the inherent attributes of self-learning, self-adaptive, associative memory and association mapping to increase the detection of ambiguity. Ref. [14] proposed a new anomaly detection algorithm, which used improved hierarchy clustering to overcome the problems of high noise and data updated. Liu et al. [15] researched user behavior of mobile terminal to mine the correlation between user pressure and user unsafe behavior to prevent the malicious and unsafe behavior of users.

In short, there are some useful researches lay solid foundations for the user dependability. It is quite important for communicating between mobile cloud and its users to complete the process of series operations in the mobile cloud environment future. Nevertheless, most researches on clustering focus on the user behaviors or softening points by the neural network model. Those researches ignore the different intrinsic property of object, while it is different in practice. The existing literatures have not yet the concrete method to calculate standardized weights and soften points completely.

3 System models

3.1 SVD model

SVD model [16] is earliest and most widely used in image processing field to reduce the time complexity and improve the efficiency during feature extraction. It has better scalability and practicality, and it is easy to integrate with other technologies to improve the performance. SVD model is as follows:

$$\begin{pmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \cdots & X_{nm} \end{pmatrix}$$

For any real matrix $\boldsymbol{A}$ with $n \times m$, there always are m orders orthogonal matrix $\boldsymbol{U}$ and n orders orthogonal matrix V, which makes $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{\Delta}=\operatorname{diag}(\delta_1,\delta_2,\ldots,\delta_r)$, $\delta_i>0(i=1,2,\ldots,r), r=\operatorname{rank}\boldsymbol{A}$, the singular values of matrix $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ are $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_r > 0$, $\lambda_{r+1} = \lambda_{r+2} = \lambda_m = 0$, we call positive number $\delta_i = \sqrt{\lambda_i}(i=1,2,\ldots,r)$ are singular value of matrix $\boldsymbol{A}$. If setting $\boldsymbol{U}=(u_1,u_2,\ldots,u_m)$, $\boldsymbol{V}=(v_1,v_2,\ldots,v_m)$, then u_i and $v_i(i=1,2,\ldots,r)$ are the eigenvectors corresponding with λ_i^2 of $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ respectively, and it introduces vector u_i and v_i is to make $\boldsymbol{U}$ and $\boldsymbol{V}$ form orthogonal matrix.

3.2 SVD de-noise model

For each information sub-matrix $X(N)=\{x_1,x_2,\ldots,$

$x_N\}$, which contains noises, constructing Hankel matrix by phase-space reconstructing.

$$R(X)_{n\times m}=\begin{pmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1m} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \cdots & X_{nm} \end{pmatrix}=D_{n\times m}+W_{n\times m} \quad (1)$$

Where $N=m+n-1$, $D_{n\times m}$ is the information sub-space without noise interfering, $W_{n\times m}$ is the noise information sub-space. The reconstructed matrix is decomposed into series of singular value with descending order. The former larger *k* singular values represent useful properties, while the latter *n-k* singular values represent noise properties. It sets the corresponding *n-k* singular values to 0 to achieve the purpose of de-noising. Then matrix R' is get by the inverse process of SVD, where R' is the optimal approximation matrix of rank $k(k<n)$ of *R*. The key of having a good de-noise result is determining the order of effective ranks and the structure of reconstruction matrix [17].

3.3 Neural network model

Neural networks [18] have characteristics of self-adaption, self-learning, self-organization, parallelism, associative memory, association mapping, better fault tolerance, higher robustness and so on. Neural networks can approximate arbitrary nonlinear function theoretically. Back-propagation (BP) neural network includes three layers, i.e. input layer, hidden layer and output layer. The number of input layer nodes is the dimensions of the feature vectors, while the number of output layer nodes is determined by the demands of user. There is no specific rule of neuron number for hidden layer. Too little neuron leads that the next layer information is not enough, which will cause the inaccurate classification result; while too much neuron leads to longer learning time and lower generalization ability. The unit number of hidden layer is based on specific cases. However, the process is time-consuming.

4 The abnormal behavior analysis mechanism

This paper analyzes the abnormal behavior of the mobile cloud users to ensure the dependability of mobile cloud environment. It researches a new method of user behavior analysis. Firstly, the user behavior dataset matrix is processed to achieve SVD and de-noise, where clustering accuracy needs to maintain. In mobile cloud environment, since data flow and mobile terminal users are large amount, map-reduce model is used to achieve parallel SVD processing to improve the decomposition speed. Then, this paper introduces BP neural network to deal with the shortcomings of hardening points in traditional clustering. In hidden layer of BP neural network, objective similarity analysis method is adopted to avoid subjective factors determine weight, where information entropy is introduced to determine and normalize the weight factor of behavior attributes. Finally, the similarity and threshold values between user behaviors and records in the normal behavior model database are calculated respectively. If the value exceeds the threshold, it is regarded as abnormal behavior. The corresponding tips and preventive measures should be made. Otherwise, it will be regarded as normal behavior and merged into the normal model database to update database in real time. The corresponding threshold will change after update the database every time. The specific framework is shown in Fig. 1.

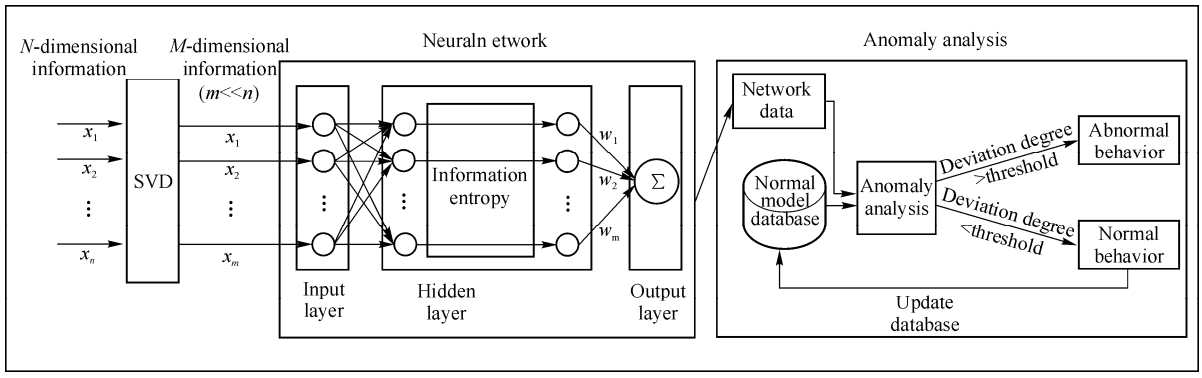

Fig. 1 Abnormal behaviour analysis framework

4.1 The parallel SVD model

Most real-world databases contain abnormal data, uncertain data, data loss and noise data. Some clustering algorithms are sensitive to such data and will lead to obtain data with poor quality. This paper uses SVD algorithm to reduce dimensionality which has higher signal-to-noise ratio (SNR), and completes the losing elements in matrix by self-learning which has a better robustness. The SVD model could screen out the useful feature information from massive amounts of information to reduce the computing complexity and improve accuracy of SNR and detection.

Serving each user as an object, behavior dataset as the attributes of object, each sub-dataset as an attribute of the object, we can construct the matrix $\begin{pmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix}$ that represents the structure of user-attribute and implements the data classification conveniently, where X_{ij} represents the j-th attribute of i-th user. Using $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^{\mathrm{T}}$ to decompose the matrix into a series of small matrix, and doing SVD for each matrix, the process results are merged into new matrixes as the input of the next iteration. SVD operation of sub-matrix is independent between each other. Thus, map-reduce parallel processing is used to accelerate the computation speed.

The kernels of the parallel compute are the two functions of map and reduce. In the map phase, attribute of each user behavior is served as a sub-task at the input port. Then, the sub-task is assigned to each server. The assigned task follows the principle of minimum transmission cost to reduce unnecessary network overhead. Each server to parallel process then serve the processing results as intermediate results temporarily stores in the local memory, where map function receives an input key/value pair and maps it as intermediate result key/value pair. In reduce phase, the intermediate result is served as an input data to merger, where reduce function deals with the intermediate results by the key value of the intermediate data to output the final result the key/value.

4.2 The SVD de-noise model

In the SVD de-noise model mentioned above, singular entropy is used to determine the number of effective orders. Firstly, we will give the concept of singular spectrum:

$$\sigma_i = \log \frac{\lambda_i}{\sum_{j=1}^{s} \lambda_j}; \quad i \leqslant s \tag{2}$$

Where $s = \min(m, n)$, the sequence composed of $\sigma_i (i = 1, 2, ..., s)$ is singular spectrum that matrix $R(X)$ by processing of singular value decomposition.

By investigating the change regularity of information followed by singular spectrum, the singular entropy can be expressed as:

$$E_k = \frac{1}{k} \sum_{i=1}^{k} \Delta E_i; \quad k \leqslant s \tag{3}$$

Where k is the order of singular entropy, ΔE_i represents the increment of singular entropy at i-th order, which could be calculated by the following formula:

$$\Delta E_i = -\left(\lambda_i \sum_{j=1}^{s} \lambda_j \right) \log \left(\lambda_i \sum_{j=1}^{s} \lambda_j \right) \tag{4}$$

When the de-noise order of the selected singular spectrum is lower, the de-noise signal contains incomplete information. Furthermore, signal waveform distortion phenomenon will occurs, which could not make accurate reflection for the effective information of the original signal. When the de-noise order of the selected singular spectrum is higher, it still remains a part of the noise information in the de-noised information, which is unable to achieve the purpose of de-noising fully. In fact, when the singular entropy increment reduce up to the asymptotic value, the effective feature information content of the signal has been reached saturation, and the feature information has been basically completed. The later singular entropy increment is led by wideband noising, which could not be considered completely. Therefore, it is reasonable that selecting singular spectrum order where singular entropy increment reduced up to the asymptotic value as the signal singular spectrum de-noising order.

Each sub-matrix after SVD processing is descending order, and descripts the importance of matrix. The higher the singular value of the column vector is, the more important the user attribute will be. It is the de-noised matrix that it removes the last n-k singular values corresponding column vector.

4.3 The neural network model

The neural network is introduced to analyze user abnormal behavior and avoid the hardening of the points. It could be identified future similar behaviors

automatically according to the historical behavior. Its association function could find the variants of the known abnormal behavior, which reduces false alarm rate (FAR) and false negative rate (FNR) of the anomaly analysis system. However, neural network cannot deal with the input with semantic form generally, and cannot ensure that the knowledge is redundant or useful. Therefore, information entropy is introduced into the hidden layer of neural network to determine the weight of each attribute and improve the deficiency of the neural network. Using information entropy to calculate weight should detect the distributed denial of service (DDOS) attacks well.

Constructing neural network model sums up roughly: firstly, in input layer, the reduced dimension matrixes are re-combined into a big matrix according to the descending singular value corresponding column vector, where the neuron number is the dimension of combined sample vector. Secondly, in hidden layer, because of non-uniformity of weights, information entropy is used to calculate the weight of each subset, where the weight should be normalized to [0, 1]. The information entropy can improve the precision of the algorithm, and it needs not consider the number of hidden layer nodes and the location and width of corresponding center node. Finally, in output layer, the accurate and normalized weight matrix $\boldsymbol{w}_{\text{norm}} = (w_{\text{norm(1)}} \quad w_{\text{norm(2)}} \quad w_{\text{norm(3)}} \quad \cdots \quad w_{\text{norm(n)}})^{\text{T}}$ is outputted, where n represents the number of users. The weights are 1 for between layer 1 and layer 2, and [0-1] for between layer 2 and layer 3 respectively.

In the neural network model, the most important step is how to measure each attribute weight in hidden layer. Information entropy is mainly used to calculate each attribute weight. The specific measures are as follows:

Eq. (5) shows how to calculate the weights of each user attributes by information entropy, e.g., for user j:

$$H\left(X^j\right) = -\sum_{i=1}^{r} p\left(x_i^j\right) \log_2 p\left(x_i^j\right) \tag{5}$$

Where x_i^j represents the i-th attribute variate of the user j, r represents attribute dimension of each user, $p(x_i^j)$ represents the probability of each user behavior emerging in the total user behaviors, $\sum_{i=1}^{r} p(x_i^j) = 1$, $0 \leqslant p(x_i^j) \leqslant 1 (i = 1, 2, ..., r)$, $j = 1, 2, \ldots, n$.

Formula (6) shows how to normalize the weight:

$$w_{norm(j)} = \frac{H\left(X^j\right)}{\sum_{j=1}^{n} H\left(X^j\right)} = \frac{\sum_{i=1}^{r} p\left(x_i^j\right) \log_2 p\left(x_i^j\right)}{\sum_{j=1}^{n} \sum_{i=1}^{r} p\left(x_i^j\right) \log_2 p\left(x_i^j\right)} \tag{6}$$

4.4 The clustering model

The traditional clustering model [19] is sensitive to noises, needs to initialize the number of clusters, and cannot analyze clusters of arbitrary shapes. Therefore, this paper proposes a novel clustering method to overcome these shortcomings during analyzing the abnormal behavior of the mobile cloud users. The precondition of analyzing abnormal behavior by clustering analysis is that there are differences between normal behavior and abnormal behavior of mobile cloud users. The similarity is calculated through comparing the current user behavior with normal behavior database. If its similarity is beyond the set threshold, it is abnormal behavior; otherwise, it is normal behavior. If the set threshold is too large, it has lower clustering accuracy and cannot achieve the effect of clustering; if too low, the clusters grow rapidly. Therefore, the core of the clustering model is the similarity calculation and the threshold set. The clustering results divide into normal behavior cluster and abnormal behavior cluster, iterating user behaviors until user is empty.

The specific method of calculating the similarity and setting the threshold is shown as follows. Defining the similarity S between the mobile cloud user a and b is

$$\begin{aligned} S &= \left(w_{\text{norm(a)}} \frac{X_i^{\text{a}}}{\| X_i^{\text{a}} \|} \right) \left(w_{\text{norm(b)}} \frac{Y_i^{\text{b}}}{\| Y_i^{\text{b}} \|} \right) = \\ & w_{\text{norm(a)}} w_{\text{norm(b)}} \left(\frac{X_i^{\text{a}}}{\| X_i^{\text{a}} \|} \frac{Y_i^{\text{b}}}{\| Y_i^{\text{b}} \|} \right) = \\ & \frac{\sum_{i=1}^{r} p\left(x_i^{\text{a}}\right) \log_2 p\left(x_i^{\text{a}}\right)}{\sum_{j=1}^{n} \sum_{i=1}^{r} p\left(x_i^j\right) \log_2 p\left(x_i^j\right)} \frac{\sum_{i=1}^{r} p\left(x_i^{\text{b}}\right) \log_2 p\left(x_i^{\text{b}}\right)}{\sum_{j=1}^{n} \sum_{i=1}^{r} p\left(x_i^j\right) \log_2 p\left(x_i^j\right)} \cdot \\ & \frac{\sum_{i=1}^{r} x_i^{\text{a}} y_i^{\text{b}}}{\sqrt{\sum_{i=1}^{r} (x_i^{\text{a}})^2 \sum_{i=1}^{r} (y_i^{\text{b}})^2}} \end{aligned} \tag{7}$$

Where n is the total number of users, and the value of Sim is between [0, 1]. The larger the value is, the greater the similarity between users is.

The threshold is set as follows:

$$\Omega = S_{\max} + (S_{\max} - S_{\min})\sqrt{\frac{S_{\max} - S_{\min}}{S_{\max} + S_{\min}}} \tag{8}$$

The steps of the algorithm are described as follows:

// The algorithm of analyzing user abnormal behavior based on neural network clustering

Step 1 The matrix that represents user behavior dataset to execute SVD.

Step 2 To determine the effective rank order according to Eqs. (2)–(4) to de-noise.

Step 3 Inputting the de-noise matrix into input layer of neural network.

Step 4 Defining the training set and testing set.

Step 5 Calculating and standardizing weight of matrix that user information of receiving in input layer by Eqs. (5), (6).

Step 6 Outputting the standardized matrix in output layer.

Step 7 Calculating the similarity between user behavior and normal behavior model database according to the Eq. (7).

Step 8 Setting the threshold of clustering according to Eq. (8).

Step 9 Comparing the similarity with threshold calculated by step 7 and 8. It is abnormal behavior that the similarity is bigger than the threshold, and the system makes the corresponding tips and precautions, otherwise it is normal behavior.

Step 10 Adding a list element of standardized weight into normal behavior and updating to the normal behavior model database for analyzing user behavior the next time.

5 Simulation experiment analysis

5.1 Experimental environment

Experimental environment: This experiment runs on ordinary PC, and the configuration follows as Table 1.

Table 1 The experimental environment

Parameters	Values
CPU	3.4 GHz Intel(R) Core(TM) i3-2130 CPU
RAM	4 GB
HDD	498GB/7200r/min
OS	Windows7
IDE	MATLAB

Related parameters: This paper uses MATLAB to simulate the algorithm and analyze the user abnormal behavior. The experimental subject is the KDD CUP99 simulation environment. We translate the proposed clustering algorithm into MATLAB frame and set the scene as in Table 2.

Table 2 Setting the simulation parameters

Parameters	Values
Sample size	6 000
Training sample	5 000
Testing sample	1 000
accuracy	0.000 1
Learning rate	0.05
Training time	100
Error change gradient	10^{-7}

Data sources KDD CUP99 dataset [20] is implemented to evaluate the proposed algorithm. It is a test dataset to evaluate intrusion detection model organized by Defense advanced research projects agency (DARPA) in Massachusetts institute of technology (MIT) Lincoln Laboratory in 1998. Later on, professor Stolfo and others at Columbia University made data preprocess furtherly for the data. Since 1999, the KDD99 dataset has widely used in evaluating abnormal detection model. KDD99 dataset contains approximately 5 million data records and each record contains 42 properties, which have been identified as normal or specific attacks. We select 6 000 data from the data as the training set randomly, where 5 000 of them as the training samples and 1 000 of them as inspection objects.

5.2 Abnormal analysis evaluation index

To detect the performance of user abnormal behavior analysis algorithm, this paper considers the five indicators, such as DS, detection rate (DR), accuracy rate (AR), FNR and FAR. A detection algorithm with fast DS, high DR, lower FAR and FNR relatively is considered as a better detection algorithm. Evaluation indexes such as DR, AR, FNR and FAR are calculated as follows:

$$I_{\mathrm{DR}} = \frac{e}{g} \tag{9}$$

Where e is the number of detected attack samples, g is the total number of attack samples.

$$I_{\mathrm{AR}} = \frac{c}{d} \tag{10}$$

Where c is the detected sample number in all abnormal samples, d is the total number of abnormal samples.

$$I_{\mathrm{FNR}} = \frac{z}{f} \tag{11}$$

Where z is the detected abnormal samples mistaken as normal samples, f is the total abnormal sample number.

$$I_{\mathrm{FAR}} = \frac{p}{t} \tag{12}$$

Where p is the number of normal behaviors mistaken as abnormal behaviors, t is the total normal samples number.

5.3 Simulation results

The massive mobile cloud service demands of users are explosive growth, if modeling directly by KDD99 dataset, it would consume large amounts of resources. It is unrealistic that each attribute of the data in the data set is useful to the performance of the algorithm. Reducing dimension successfully can greatly reduce modeling time, improve the computing speed and maintain the clustering accuracy. The following simulation compares the five indicators such as DS, DR, AR, FNR and FAR respectively, where DR, AR and FNR are complementary between each other. All of the following simulation results are the average of ten random simulation experiment results.

Fig. 2 shows the DS comparison among user abnormal behavior analysis method based on clustering (UABAC), user abnormal behavior analysis method based on neural network (UABANN) and user abnormal behavior analysis method based on neural network clustering (UABANNC). As seen in Fig. 2, the DS of UABANNC is not as good as other two algorithms when number of samples is relatively less. It is because that the proposed algorithm needs to reduce dimensions and de-noise by SVD in the early stage. It presents the basic consistency when the samples reach 80. After that, with the increasing of testing samples, the test speed has a significantly improve because of the dimension reduction data becomes more simple and clear than the original data.

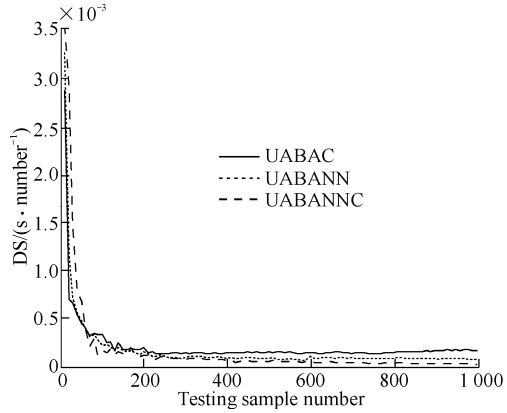

Fig. 2 Comparison of DS

An algorithm with higher DR can analyze the abnormal behavior more accuracy and interrupt the smooth progress of aggressive behavior to protect personal data of users effectively. The Fig. 3 shows that DR can reach 100% while test sample is fewer. DDOS attack is occurred when testing samples reach at 30~90. UABAC cannot detect the attack behavior well and mistakenly think the behavior is normal behavior, so the DR is rapidly decreased. In this case, the algorithms of UABANN and UABANNC could detect the attacks well, so the DR remains stable. With the increasing of testing sample, UABANNC shows obvious advantages than the other two algorithms. This algorithm eliminates noise and irrelevant attributes interference of the data, which makes the DR is higher and more stable.

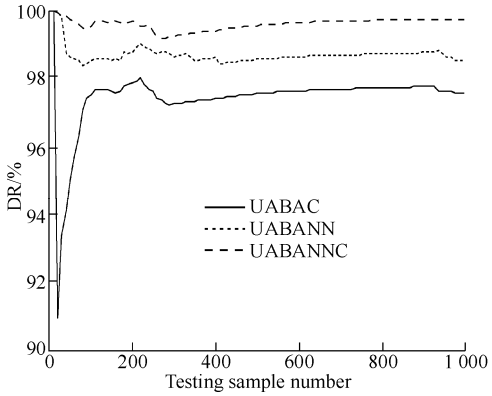

Fig. 3 Comparison of DR

AR is the ratio of detected abnormal samples and all the abnormal samples. It can be seen from Fig. 4, AR of the three kinds of user abnormal analysis algorithms can reach 100% while testing samples are fewer. AR reduces sharply when the number of samples is about 30~90, which in that it occurs DDOS attacks. The other two kinds of algorithms could detect the DDOS attacks to make the AR stable relatively. With the increase of sample size, UABANNC has a higher DR compared with the other two algorithms, which because that the algorithm reduces some data noises and irrelevant attributes interferences.

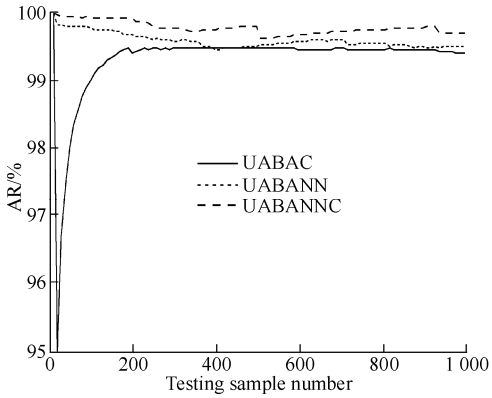

Fig. 4 Comparison of AR

FNR reflects the recognition capabilities for user abnormal behavior. The ideal situation is that the lower the FNR is, the better the algorithm will be. It can be seen from Fig. 5, the FNA of the three algorithms are almost same while testing samples are fewer. DDOS attacks are occurred when the number of samples is about 30~90. UABAC cannot identify DDOS attacks, so FNA is higher than the other two algorithms. With the increasing of testing samples, UABANNC has little effected by noise than the other two algorithms, so it has lower FNA. it shows that this method has better scalability, self-adaptability and higher recognition ability.

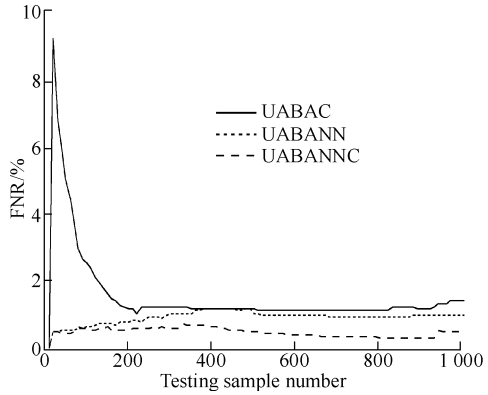

Fig. 5 Comparison of FNR

FAR reflects the ability of identifying the normal samples. If FAR is very high, the real hazard warning information would be submerged in the useless alarm information, which makes the abnormal behavior is executing successfully. It can be seen from Fig. 6, with the increasing of testing samples, FAR of the three kinds of algorithms are increasing gradually. UABANNC has lower FAR than the other two algorithms. It shows that the algorithm could better recognize the user abnormal behavior.

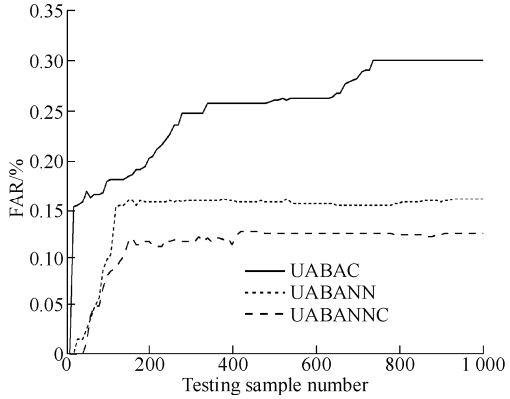

Fig. 6 Comparison of FAR

The experimental results show that the proposed algorithm improves the detection accuracy of the user abnormal behavior and ensures the DS, and improves the FAR and FNA too. From the above simulation results, we can see that the algorithm has a good stability. It can identify user abnormal behavior effectively and has a better scalability and self-adaptability. After several experimental verifications, the simulation results in accordance with above results.

6 Conclusions

The mobile Internet business is bound to explosive growth in big data era, and all kinds of intrusion behaviors make information security face severe tests. From the aspects of user dependability, this paper proposes a user abnormal behavior analysis method based on neural network clustering. Firstly, map-reduce model is adopted to execute SVD, which makes for massive data to reduce dimensions and de-noise. Secondly, neural network for softening points is used to solve over-fitting phenomenon during clustering. Finally, information entropy is used in hidden layer of neural network model to calculate the weight of each attribute and solve the problem of flooding feature attribute caused by equal weight of attributes. Both analytical and simulation results indicate that our scheme outperforms other schemes to a certain extent. Our scheme could better establish the mutual trust relationship between users and clouds. However, this method could only analyze abnormal behavior, but not for further processing. In future work, we should deep scan abnormal behavior to predict service steps, service effect and the scope of the diffusion effect, which may affected by abnormal time sequence behavior.

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (U1404611, U1204614, 61370221), in part by Program for Science& Technology Innovative Research Team in University of Henan Province (14IRTSTHN021), and in part by the Program for Science & Technology Innovation Talents in the University of Henan Province (14HASTIT045, 16HASTIT035), in part by Henan science and technology innovation outstanding talent (164200510007).

To p. 44