Send Orders for Reprints to reprints@benthamscience.ae

1482 *The Open Automation and Control Systems Journal*, 2015, *7*, 1482-1485

Open Access

Application of Clonal Selection Clustering Algorithm for Anomaly Detection in Network Security Management

Qian Zhang[1,*], Xiaoyu Wang[2] and Yan Li[1]

[1]*Department of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, Jiangsu, China;* [2]*The Attached Middle School of Xizang Minzu University, Xianyang 712082, Shaanxin, China*

Abstract: Network security being of practical significance, the importance of network application, the network anomaly detection and the generalization ability are studied as a key link in the network security management. The key technology of network security management is based on: artificial intelligence theory as the research object; the clonal selection method of anomaly detection based on fuzzy clustering algorithm, to solve the anomaly detection of low efficiency; high false alarm rate, proposes the compensation method of evidence combination rule based on the rule of sharing, to solve the problem of the information fusion of evidence combination rule of conflict and defects; and P2P trust management model based on the improved evidence combination rule to solve the P2P system. It is however, difficult to effectively deal with the malicious node attack, besides it can't effectively deal with the uncertain information and other issues.

Keywords: Bending, distortion, Matlab, tilt, the center point.

1. INTRODUCTION

Internet application involves all fields of society, such as electronic bank, electronic commerce, social networking and others. Network security management is an important part of the protection of personal privacy and account security. It is the future development of network application that is one of the important topics [1-5]. Network security management mainly includes intrusion detection and trust management, that is achieved using artificial intelligence, the clustering theory and evidence theory to carry out the management of the network security, which is a clonal selection theory of clustering based on fuzzy clustering algorithm; the clonal selection method of anomaly detection based on fuzzy clustering algorithm for Intrusion Detection based on evidence theory, which proposes compensation method of the evidence combination rule and reliability of the average compensation based on the sharing of evidence combination rule based trust management; and trust management model for P2P based on the improved evidence combination rule [6-10]. For large data sets, on one hand, the number of samples is large, this is the clustering algorithm especially the classical clustering algorithms can be challenging; on the other hand, increasing the number of categories of data sets, will result in some categories of the samples within a class or less. The within class distribution is relatively dense, and exists in the sample space approximately in isolated form. The data set for clustering algorithm especially the clustering algorithm based on natural computation, clustering the data demands higher requirements. In the existing clustering algorithms to cluster the data, the clustering center during the iteration process is easy to fall into local optimal value, to obtain the correct clustering of the data set. Therefore, the clustering analysis of large-scale multi class data established that clustering algorithm needs to improve in reliability and stability.

Recently, artificial immune system algorithm has been successfully applied for clustering analysis of clone selection, to further improve the performance of clustering algorithm. It is based on artificial immune system, which is suitable for clustering the large scale, multi category complex data sets. This paper thus proposes a new clustering algorithm, *i.e.* the framework of the selection algorithm based on clone into immunodominance operator, the algorithm can accumulate prior knowledge of the search process on the clustering centers, speed up the convergence process, to prevent falling into local optimum, and ensure the stability of the high extraction algorithm.

Thus, the key technology of network management, i.e. the theory of artificial intelligence will focus on intrusion detection and trust management.

2. CLONAL SELECTION METHOD OF ANOMALY DETECTION BASED ON FUZZY CLUSTERING ALGORITHM

In the management of network security, intrusion detection is to collect and analyze the network application in the process of information transmission, such as the discovery of malicious or breach of security strategy, thus it allows to take defensive measures [11]. The clonal selection method of anomaly detection algorithm of fuzzy clustering based on the strong intrusion ability to detect unknown attacks effectively, can control and improve the detection rate and reduce false positive rate.

2.1. Cluster Analysis

Cluster analysis is to group information network on the basis of the maximum and minimum within class similarity

*Address correspondence to this author at the Department of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, Jiangsu, China; E-mail: qianzhangqzqz@126.com

1874-4443/15

and between class similarity. A cluster analysis of data model, set $X=\{x_1,x_2...,x_n\}$ as a sample set, where $x_i=\{x_{i1},x_{i2}...,x_{im}\}$ for the I samples the M characteristic values, $X_{ij}=\{j=1,2,..m\}$ for the j feature assignment, $P(x_i)$ as the feature vector of x_i. The cluster analysis is a partition of sample sets into several disjoint subsets, a partition number is k, M is a fuzzy control factor, the fuzzy clustering objective function is as follows:

$$C(U,P)=\sum_{i=1}^{k}\sum_{l=1}^{n}\mu_{il}^{m}\left[D(x_i,p_i)\right]^2,\mu_{il}\in[0,1] \quad (1)$$

Cluster analysis for the sample characteristic and sample concentration can construct a hierarchical, classification, probability density, through the clustering method based on grid model. Hierarchical clustering method is a hierarchical decomposition of the sample set into subset; clustering method is constructed by dividing the data set; probability density clustering method is a neighborhood density threshold set for the data set clustering according to the threshold; grid clustering method is to quantify the sample set into grid structure, according to the unit data clustering; clustering method based on the assumed model is established, according to the clustering samples, data and model [12].

2.2. Clonal Selection

In 1958, Burnet *et al.* proposed immune clonal selection theory. The famous clonal selection principle is that, in addition to amplify lymphocyte differentiation into plasma cells, it can differentiate into B memory cell life cycle longer. When encountering the corresponding antigen, memory cells are preselected by the immune system for rapid activation, proliferation, differentiate into antibody producing cells, for the implementation of efficient and durable immune function.

Inspired by the clonal selection theory, De Castro *et al.* proposed a clonal selection algorithm (CSA), antibody clonal selection mechanism with the help of the immune system, and clonal structure suitable for artificial intelligence. Based on the cloning operator selection algorithm is a population search strategy, ensuring a parallel random and search change, easy to fall into local optimal value in the search, and that global optimum can with high probability obtain the solution of the problem, and the convergence speed. Based on the above advantages, CAS is used as the main frame structure of the new clustering method.

Immunology model suggests, although an antigen molecule can have multiple epitopes, yet in the induction of host immune response only one or a few epitopes may play a major role in the host immune response to the specific target; this phenomenon is called immune advantage or immune dominance. The key agent is known as a dominant epitope. Immune advantage is produced due to the interaction between antibody and antigen, and it is a dynamic process. The immune dominant site is decided by natural selection in which an antigen will face greater pressure.

In the artificial immune system, immune dominance was programmed to represent each of the antigen encoding different antibodies at certain important degree. It needs to be pointed out that the biologically defined artificial system does not completely follow the immune dominance, partly because it is the basis of idiotypic regulatory network theory, and has the characteristics of antibody antigen, so the concept of antibody has a biological basis. On the other hand, the main operation of artificial immune system depends on antibody, and antigen.

Clonal selection is set up on the basis of the biology of the immune system in a clonal selection algorithm, which is the algorithm based on evolutionary search *vs*. random search, local search vs. global search characteristics to create a clone operator, since the convergence speed of the clone operator is higher [13]. Clonal selection algorithm and evolutionary algorithm, assuming the antigen function, are as follows:

$$\phi:\prod_{i=1}^{m}[d_i,u_i]\to R(d_i<u_i),x_i\in[d_i,u_i],$$

M is the number of optimization variables, antibody of A in the N group $\overline{A}=\{A_1,A_2...,A_n\}$, antigen affinity function is defined as:

$$W_{ij}=\left\|A_i-A_j\right\|,i,j=1,2,...,N \quad (2)$$

2.3. Fuzzy Clustering Algorithm Based on Clonal Selection

Fuzzy clustering algorithm is shown in Fig. (**1**) based on clonal selection.

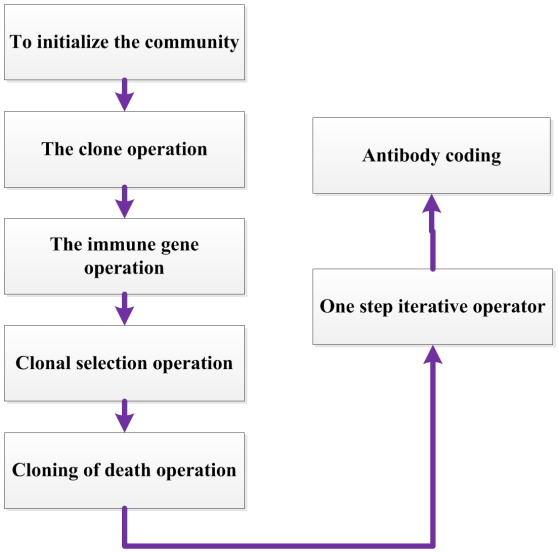

Fig. (1). Fuzzy clustering algorithm based on the clonal selection process.

To initialize the randomly generated community $A(0)=\{A_1(0),A_2(0)...,A_K(0)\}$; calculation of affinity $\{J(A(0))\}$; $k=0$; to determine whether they meet the iteration termination conditions; implementation of the clone operation of A'(k); implementation of the immune gene operation A"(k); implementation of clonal selection operation to get new antibody in group A(k+1); cloning operation to meet death

$$f(A_i(k+1))=f(A_j(k+1))=\max\{f(A(k+1))\},i\neq j,$$

then the affinity; k++ returns to the iteration termination judgment.

3. TO MAKE THE RULE OF EVIDENCE COMBINATION BASED ON SHARED PRINCIPLES

Network security management is essential to obtain evidence. But the evidence is diverse and there is uncertain information fusion, so in order to be able to more effectively collect information, it is proposed to make evidence sharing combination rule theory based on problem solving, conflict and limited data information [14].

3.1. The Rules of Thinking

Establishment of evidence information collection of X and Y, drawing set as shown in Fig. (**2**).

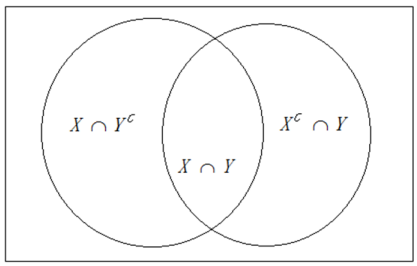

Fig. (2). Venn diagram set X and Y.

Set X and Y to make the relationship in 4 ways: $X^C \cap Y^C$, $X^C \cap Y$, $X \cap Y^C$, $X \cap Y$. Using a binary 0, 1 expressed X, Y complement and original set, $X^C \cap Y$ expressed as $(01)_2$, $X \cap Y^C$ expressed as $(10)_2$, $X \cap Y$ expressed as $(11)_2$, $X^C \cap Y^C$ for external parts shall not be assigned.

3.2. Proportional/Share the Evidence Combination Rule

Set the focal element X and Y, E_1 and E_2 in the proof of the existence of information collection of X, Y BBA can be expressed as $m_1(X)$ and $m_2(Y)$.

$$\begin{aligned} &m_1(X) = n_1(XY) + n_1(X^Y); n_1(XY) = \lambda_1 m_1(X), \\ &n_1(X^Y) = (1-\lambda_1) m_1(X); \qquad \lambda_1, \lambda_2 \in [0,1] \end{aligned} \tag{3}$$

$$\begin{aligned} &m_2(Y) = n_2(XY) + n_2(Y^X); n_2(XY) = \lambda_2 m_2(X), \\ &n_2(Y^X) = (1-\lambda_2) m_2(X); \end{aligned} \tag{4}$$

Combination of $m_1(X)$ and $m_2(Y)$ BBA is as follows:

$$\begin{aligned} &m_1(X) m_2(Y) = \{n_1(X^Y) + n_1(Y^X)\}\{n_2(X^Y) + n_2(Y^X)\} = \\ &M(X^Y) + M(X^Y) + m(Y^X) \end{aligned} \tag{5}$$

$$m(XY) = n_1(XY) n_2(XY) + \frac{\{n_1(XY)\}^2 n_2(Y^X)}{n_1(XY) + n_2(Y^X)} + \frac{n_1(X^Y)\{n_2(Y^X)\}^2}{n_1(X^Y) + n_2(XY)} \tag{6}$$

$$m(X^Y) = \frac{\{n_1(X^Y)\}^2 n_2(XY)}{n_1(X^Y) + n_2(XY)} + \frac{\{n_1(X^Y)\}^2 n_2(Y^X)}{n_1(X^Y) + n_2(Y^X)} \tag{7}$$

$$m(Y^X) = \frac{\{n_2(Y^X)\}^2 n_1(XY)}{n_1(XY) + n_2(Y^X)} + \frac{\{n_2(Y^X)\}^2 n_1(X^Y)}{n_1(X^Y) + n_2(Y^X)} \tag{8}$$

3.3. The Average Pay up to Share the Reliability of Evidence Combination Rule

The average reliability of evidence combination rule and the compensation allocation proportion to make sharing evidence combination rule, is different parameters selection method of λ_1, λ_2, to fill the share proportion of evidence combination rule, $\lambda_1, \lambda_2 \in [0,1]$, and the reliability of evidence combination rule to make the average share in $XY = \varphi, \lambda_1 = \lambda_2 = 0$; if $X^Y = \varphi$, $\lambda_1 = 1$, $Y^X = \varphi$, $\lambda_2 = 1$; on the condition of $\lambda_1 = \lambda_2 = 0.5$.

4. THE P2P TRUST MANAGEMENT MODEL BASED ON THE IMPROVED EVIDENCE COMBINATION RULE

4.1. Modeling of BBA Node

Network data information marked by the trust documents and malicious files are defined as G (good) and M (Malicious) respectively, the definition of frame of discernment. A node of the I node of J recommendation reputation information confidence level is α_{i1}, α_{i2}, $\alpha_{i1} + \alpha_{i2} = 1$, of any node R_i in the $\theta_j \in \{G, M\}$, j=1,2 support strength of $m(\{\theta_j\} | R_i) = \lambda \alpha_{ij} = 1,2; \lambda = (0,1]$;

$$\sum_{j=1}^{2} m(\{\theta_j\} | R_i) = m(\{G\} | R_i) + m(\{M\} | R_i) \le 1 \tag{9}$$

Recommended that the nodes j and n, can produce n and j interactions of {G}, {M} supporting the recommendation evidence.

4.2. Evidence of Pre Treatment

Evidence handling is carried out through the consistency test between the evidence and the evidence of a source of evidence. To determine the credibility of evidence, firstly the possibility of each focal attribute is confirmed; secondly, through the collection of evidence supporting focal element of possibilities, the credibility of evidence is judged. The greater the possibility of focal elements support, the bigger the credibility of evidence, and the credibility of the minimum. The credibility of evidence is reflected through the consistency between the evidence and the evidence of other evidence, which can set a threshold, as the demarcation point of credibility, i.e. if credibility is greater than the threshold it is regarded as credible evidence, if less than the threshold it is regarded as disturbance of the credibility of evidence [15].

4.3. P2P Trust Model

Based on the process improvement of combination rule of evidence, P2P trust model is shown in Fig. (**3**).

First, determine the credible degree recognition framework of trust file G and malicious files under M; secondly, by calculating the basic belief assignment function information recommendation, the preprocessing method of interference information filtering algorithm, is used to make each evidence independent of each other; third, undertake the average pay compensation to share evidence reliability combination rule fusion recommendation information; according to the degree of trust evaluation method to determine the node trust.

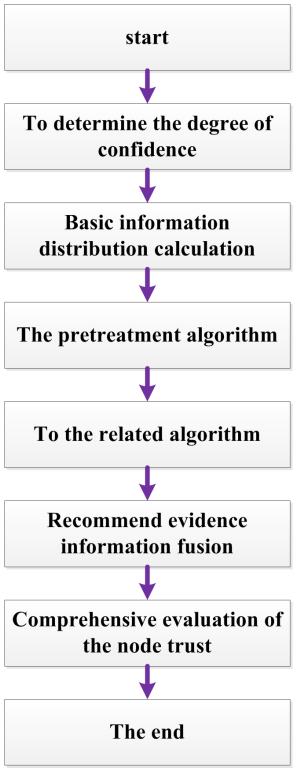

Fig. (3). P2P trust model operation process.

CONCLUSION

In this paper the intrusion detection in network security management and information management research, through the study of network security management was studied based on the theory of artificial intelligence technology. The theory was proposed based on the proportion to make sharing rule combination rule of evidence theory based on reliability and average pay compensation. Sharing rule combination rule of evidence, effectively solved the difficult problems of development and research of theory of the evidence.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] J. Liu, W. Xie, J. Huang, and W. Li, "Clustering analysis by genetic algorithms," *Acta Electronica Sinica*, vol. 23, no. 11, pp. 81-83, 1995.

[2] L.O. Hall, I.B. Ozyurt, and J.C. Bezdek, "Clustering with a genetically optimized approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 7, pp. 103-112, 1999.

[3] M. Ujjwal, and B. Sanghamitra, "Genetic algorithm-based clustering technique," *Pattern Recognition*, vol. 33, no. 9, pp. 1455-1465, 2000.

[4] L. Tseng, and S. Yang, "A genetic approach to the automatic clustering problem," *Pattern Recognition*, vol. 34, no. 2, pp. 415-424, 2001.

[5] M. Cowgill, R. Harvey, and L. Watson, "A Genetic algorithm approach to cluster analysis," *Computers and Mathematics with Applications*, vol. 37, no. 7, pp. 99 -108, 1999.

[6] J. Liu, and W. Zhong, "A novel clustering based on the immune evolutionary algorithm," *Acta Electronica Sinica*, vol. 29, no. 12, pp. 1868-1872, 2001.

[7] X. Xing, J. Pan, and L. Jiao, "A novel K-means clustering based on the immune programming algorithm," *Chinese Journal of Compuetrs*, vol. 26, no. 5, pp. 605-610, 2003.

[8] J. Li, X. Gao, and L. Jiao, "A CSA-based clustering algorithm for large data sets with mixed numeric and categorical values,' *Acta Electronica Sinica*, vol. 32, no. 3, pp. 357 -362, 2004.

[9] M. Gong, L. Jiao, W. Ma, and X. Zhang, "Unsupervised classification and recognition using an artificial immune system based on manifold distance," *Acta Automatic Sinica*, vol. 34, no. 3, pp. 367 -375, 2008.

[10] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems," *Complex Systems*, vol. 13, no. 2, pp. 87-129, 2001.

[11] Li Jie, Gao Xin-bo, Jiao Li-cheng, A CSA-based new fuzzyclustering algorithm. Journal of Electronics & Information Technology. 2005, 27(2): 302 -305.

[12] M. F. Burnet, "A modification of Jernes theory of antibody production using the concept of clonal selection," *Austrian Journal of Science*, vol. 20, no. 1, pp. 67-76, 1957.

[13] L. N. De Castro, and F. J. Von Zuben, "Artificial Immune Systems: Part I-Basic Theory and Applications," *RT DCA*, Brazil, vol. 95, 1999.

[14] H. Du, L. Jiao, and R. Liu, "Immunodomaince clone algorithm," *Journal of Electronics & Information Technology*, vol. 26, no. 12, pp. 1918-1924, 2004.

[15] U. Bandyopadhyay, and U. Maulik, "Nonparametric genetic clustering: comparison of validity indices," *IEEE Trans. Syst, Man, Cybern. C*., vol. 31, no. 1, pp. 120-125, 2001.

Received: May 26, 2015 Revised: July 14, 2015 Accepted: August 10, 2015