Application of Student Real Name Data Information Processing Based on Artificial Intelligence Technology in Network Security

Wenyan Ye
School of Marxism
Hohai University
Nanjing, Jiangsu, 211100, China
1063540541@qq.com

***Abstract*—With the continuous development of network technology as well as science and technology, artificial intelligence technology and its related scientific and technological applications, in this process, were born. Among them, artificial intelligence technology has been widely used in information detection as well as data processing, and has remained one of the current hot research topics. Those research on artificial intelligence, recently, has focused on the application of network security processing of data as well as fault diagnosis and anomaly detection. This paper analyzes, aiming at the network security detection of students' real name data, the relevant artificial intelligence technology and builds the model. In this process, this paper firstly introduces and analyzes some shortcomings of clustering algorithm as well as mean algorithm, and then proposes a cloning algorithm to obtain the global optimal solution. This paper, on this basis, constructs a network security model of student real name data information processing based on trust principle and trust model.**

Keywords*—*Artificial intelligence, Network security, Student real name data, Data processing

I. Introduction

A. Overview of Student Data Network Security Management Based on Artificial Intelligence Technology

Artificial intelligence technology, as a frontier branch of computer science, has attracted extensive attention in the context of Internet big data. The main research direction of artificial intelligence technology in China, recently, proves to combine artificial intelligence with network security to effectively manage network information as well as prevent information loss. The main technologies, in this process, are made use of include machine learning, data mining, computer technology, computer management system and network server system.

The main research direction, nowadays, in this field, keeps the combination of data clustering as well as network security. Therefore, this paper, in essence, is mainly committed to the research on the application of artificial intelligence in network security, whose purpose proves to find a technology to combine artificial intelligence as well as data clustering in a strikingly effective manner.

Intrusion detection system, as the most commonly utilized protection system for network security, can identify and intercept internal and external attacks, so as to protect the system in real time. In addition, the above methods, generally, are called misuse detection methods. In this way, by collecting as well as counting the relevant data in the process of network security, we can find the abnormal behavior of the system to detect the situation of network security and system attack. Moreover, as the main direction of domestic research, at present, intrusion detection system usually utilizes the relevant data between the established normal server as well as users to formulate an abnormal activity data table, whose purpose keeps to judge the attack of the system in time and effectively.

B. Overview of Abnormal Data Detection Algorithms Based on Fuzzy Clustering

So far, the detection method of data entry detection for intrusion system, generally, is usually made advantages of. In addition, this detection method keeps inefficient, proves prone to false positives, and has less adaptable. Moreover, it is not suitable for the current network with large amount of data. Consequently, a new network security system is needed, for effective management and students' real name data prevention [1].

Clustering algorithm, in a general manner, is made use for multivariate analysis. This method of statistics effectively classifies the data via looking for the similarity between the data to form one adjacent digital data set after another. In addition, this kind of learning methods usually utilize the observation learning method to learn and train the data samples, which can deduce the corresponding rules in the data set, classify the clustering problem and turn it into an optimization issues [2].

Generally, cluster analysis uses pattern analysis to classify the data, and this classification includes two types, one proves supervised, and the other is unsupervised. For one thing, supervised analysis usually analyzes the known data to analyze all the attributes of the data. For another, the unsupervised analysis method usually uses the clustering method to divide the data set into clusters, so as to identify the degree of closeness between them [3]. All in all, the advantage of this classification method is that it does not need prior knowledge, so it needs less learning samples and has a wide range of applications. At the same time, several methods of fuzzy clustering and their comparison relations are illustrated in the Figure 1 below.

978-1-6654-0034-3/22/$31.00 ©2022 IEEE
DOI 10.1109/ACCTCS53867.2022.00071

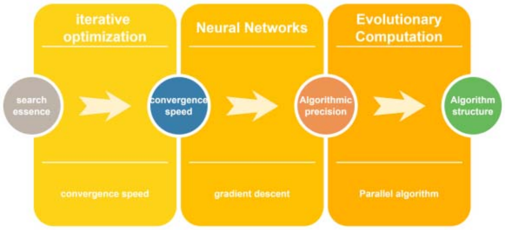

Figure 1. Relationship and comparison of fuzzy clustering methods

C. Issues of Abnormal Data Detection Algorithm Based on Fuzzy Clustering

The clustering method utilized, at present, usually carries out hard division for the data, and this hard division method has relatively high algorithm efficiency, however, its classification method proves relatively rigid and can not divide the data in a flexible manner, which is going lead to data classification errors [4]. The fuzzy clustering method used, so far, can produce effective connections between data categories. Moreover, the optimization effect of this clustering method keeps better than the above methods, but its disadvantages never fail to be outstandingly obvious as well. Some data sets cross in the clustering process, and it keeps difficult to distinguish the crossed data sets, so it needs to be further optimized.

II. Clustering Model of Student Real Name Data Information Processing

A. Cluster Analysis Process

In the process of data set classification, firstly, the authors define the sample for the data sample. Moreover, in the process of defining the sample set, the corresponding assignment is mainly carried out through its characteristics. The assignment process mainly uses eigenvalues to represent the position of this data in its corresponding matrix to mark the corresponding subset, and the data eigenvalue identification process is as follows.

$$X_1 \cup X_2 \cup \cdots \cup X_k = X; X_i \cap X_j = \varnothing, 1 \le i \ne j \le k \tag{1}$$

After data classification, the disjoint subsets are divided into membership relations through cluster analysis, so as to establish mapping. The process is as follows.

$$\mu_{X_i}(x_l) = \mu_{il} = \begin{cases} 1, & x_l \in X_i \\ 0, & x_l \notin X_i \end{cases} \tag{2}$$

After the mapping relationship is established, the subsets, in the membership function, are non-empty, and the non-empty subsets are identified. In addition, generally, this method is called hard partition, which in the clustering process, proves the first step. The unit vector, in the division process, is defined for the matrix in the unit data, so as to select its vector direction. The value method is as follows.

$$M_h = \left\{ \mu_{il} \mid \mu_{il} \in \{0,1\}; \sum_{j=1}^{k} \mu_{jl} = 1, \forall l; 0 < \sum_{j=1}^{n} \mu_{jl} < n, \forall i \right\} \tag{3}$$

In the process of fuzzy clustering, the student real name data is divided into several samples by dividing the sample set. These sample sets have subordinate relations with each other, and their subordinate relations are shown in the above formula [5]. Moreover, these data sets proves of specific significance, and divided according to appropriate definitions. For the measurement of these data sets, generally, the Euclidean distance between them is used to express the degree of data reciprocity. The specific representation is as follows.

$$M_h = \left\{ \mu_j \in M_h \mid \mu_{ji} \in \{0,1\}, \forall i \right\} \tag{4}$$

In order to ensure the reliability and accuracy of clustering algorithm, the divided data are usually processed by minimizing data. This processing method can calculate and construct the distortion of data samples, so as to process the function in the fuzzy region. On this basis, the authors optimize the objective function, so that the optimization algorithm can act on the clustering mode. The optimal clustering optimization method is as follows.

$$M_f = \left\{ \mu_{il} \mid \mu_{il} \in [0,1]; \sum_{j=1}^{k} \mu_{jl} = 1, \forall l; 0 < \sum_{j=1}^{n} \mu_{jl} < n, \forall i \right\} \tag{5}$$

After the above calculation process, the data can be divided into subsets. In addition, the data set, on this basis, can be further analyzed and processed in detail.

B. The Selection Process of Student Real Name Data Cloning

When clonal data algorithm was first proposed, it was used for group search. In the later application process, due to the selection of cloning operator and through the group optimization strategy, parallel algorithm processing and search range processing, in essence, can also be carried out, so it is also used for the calculation of global optimization algorithm [6].

Therefore, in the research process of this paper, the clonal selection algorithm is applied to network security, so as to ensure that it acts as a supplement to the clustering algorithm in the process of network security processing of student real name data, so that it can maintain stable operation in the process of large-capacity data input. The specific process is shown in the Figure 2 below.

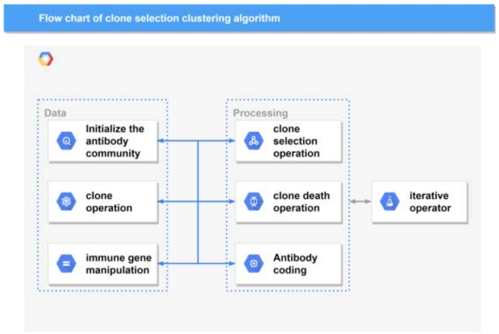
Flow chart of clone selection clustering algorithm
Data
Initialize the antibody community
clone operation
immune gene manipulation
Processing
clone selection operation
clone death operation
Antibody coding
iterative operator

Figure 2. Flow chart of clonal selection fuzzy clustering algorithm

The specific process of cloning algorithm is demonstrated above. To begin with, by dividing the antibody community of the data set, it can calculate the initial affinity. Moreover, the different affinity is used to cluster the data sets. In addition, on this basis, the clone scale of antibody is set, and the clone operator is calculated to obtain the next generation population data set [7]. Finally, on this basis, the probability sequence of the clone operator is calculated by calculating the Markov chain. The process is illustrated above.

C. Establishment of Fuzzy Clustering Algorithm Model Based on Clonal Selection

Through the analysis of the cloning algorithm in the previous section and its application in the network security model of student real name data processing, the specific implementation process of fuzzy clustering algorithm based on clonal selection can be obtained. In addition, the establishment process of clonal selection fuzzy algorithm proves mainly by decoding the clustering problem into the antibody problem [8]. Moreover, the antibody affinity function is measured and constructed, so that the items of clustering problem can be characterized [9]. In this process, it proves more than indispensable for how to select the clustering operator and how to select the operation parameters.

III. Network Security Model of Student Real Name Data Information Processing Based on Amelirated Evidence Combination

A. Establishment of Evidence Combination Rule Model

In the process of establishing evidence selection model, it is mainly made use of to define the approximation algorithm for the data. In this process, the purpose proves to generate effective information reasoning when the data proves uncertain [10]. Generally, this algorithm is used for the fusion as well as generalization of evidence and information data.

B. Establishment of Trust Model

In the process of evidence data synthesis, a certain amount of error data will be generated. Therefore, in the process of evidence theory, credibility analysis is usually carried out for the data analyzed by people. On this basis, a set of possible results is set to analyze and represent the trust model.

First and foremost, the trust function is constructed and defined for the students' real name data information. Through the identification framework, the domain is defined, so as to define an effective trust function, as demonstrated below.

$$m(XY) = \frac{1}{2} m_1(X) m_2(Y) \tag{6}$$

After the above definition, the basic reliability of compatibility in the data, by defining and calculating the natural function, can be calculated, as demonstrated below.

$$m\left(X^Y\right) = \frac{1}{2} \frac{m_1(X)}{m_1(X)+m_2(Y)} m_1(X) m_2(Y) = \frac{1}{2} \frac{\{m_1(X)\}^2 m_2(Y)}{m_1(X)+m_2(Y)} \tag{7}$$

The natural function and reliability function have the following relationship in the calculation process.

$$m\left(Y^X\right) = \frac{1}{2} \frac{m_2(Y)}{m_1(X)+m_2(Y)} m_1(X) m_2(Y) = \frac{1}{2} \frac{m_1(X)\{m_2(Y)\}^2}{m_1(X)+m_2(Y)} \tag{8}$$

In the calculation process, generally, the calculation of trust interval is involved. The measurement of trust interval is demonstrated in the following formula.

$$m(XY) = \frac{1}{2} m_1(X) m_2(Y) + \frac{m_1(X)\{m_2(Y)\}^2}{m_1(X)+2m_2(Y)} \tag{9}$$

The trust interval in the data set is usually given by the important information of the proposition. In the given process, the eigenvalues of the data set are involved. The naming process is illustrated below.

$$m\left(X^Y\right) = \frac{1}{2} \frac{\{m_1(X)\}^2 m_2(Y)}{m_1(X)+2m_2(Y)} \tag{10}$$

The data, in the process of evidence combination, is usually represented in terms of trust. In the process of representation, the data weakness is found through conflict evidence, so as to judge the network security. The definition process is illustrated in the following formula.

$$m\left(Y^X\right) = 0 \tag{11}$$

$$m(XY) = 0 \tag{12}$$

In this process, generally, the trust data is represented by combination rules. The trust degree of combination rules is calculated as follows.

$$m\left(X^Y\right) = \frac{m_1(X)}{m_1(X)+m_2(Y)} m_1(X) m_2(Y) = \frac{\{m_1(X)\}^2 m_2(Y)}{m_1(X)+m_2(Y)} \tag{13}$$

$$m\left(Y^X\right) = \frac{m_2(Y)}{m_1(X)+m_2(Y)} m_1(X) m_2(Y) = \frac{m_1(X)\{m_2(Y)\}^2}{m_1(X)+m_2(Y)} \tag{14}$$

C. Network Security Trust Model Structure Based on Improved Evidence Combination

The ameliorated network security trust model is mainly developed and simulated through P2P network. In addition, the main way of this network keeps to analyze the trust degree of data. In this process, the trust management mechanism is used to classify and judge the data, so as to combine network services with information security processing. The specific process is demonstrated in the Figure 3 below.

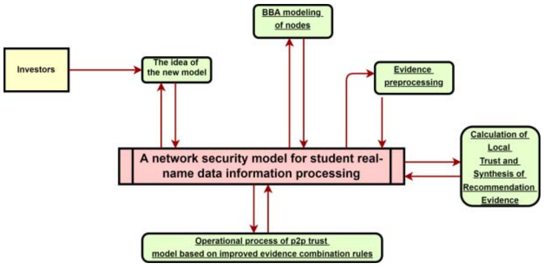
BBA modeling of nodes
Investors
The idea of the new model
Evidence preprocessing
Calculation of Local Trust and Synthesis of Recommendation Evidence
A network security model for student real-name data information processing
Operational process of p2p trust model based on improved evidence combination rules

Figure 3. Network security model framework of student real name data information processing based on improved evidence combination

This paper mainly studies artificial intelligence technology. On this basis, through the analysis of the application of artificial intelligence technology in the process of cluster analysis, this paper obtains how to cluster students' real name data. On this basis, this paper analyzes the advantages and disadvantages of the above methods, points out their related shortcomings, and additionally, puts forward the deduction method of data cloning algorithm. After the above analysis, this paper constructs the relevant content of the artificial intelligence technology student network security real name data information model. This paper, in this way, puts forward some prospects for this direction, which provides some research ideas for other researchers, and the future content of this direction is mainly network security.

With the continuous upgrading and expansion of Internet technology, cyberspace's development has become very rich in its components. The emergence of artificial intelligence has made cyberspace more flexible, and the spatial structure has developed in a more intelligent and complex direction. Combined with artificial intelligence technology's simulation of human thinking and its strong learning ability for big data, from the application results of artificial intelligence technology in cyberspace security defense, the flexible combination of artificial intelligence technology and security defense technology can be better. To make up for the shortcomings of traditional cyberspace security defense. Artificial intelligence is the product of social development and technological innovation and an important technological form that promotes human progress. Since the development of artificial intelligence, it has become the core driving force of a new round of scientific and technological revolution and industrial transformation and is having a profound impact on the world economy, social progress, and people's lives. In the world economy, artificial intelligence is a strategic technology leading the future. Major countries and regions in the world regard the development of artificial intelligence as a major strategy to enhance national competitiveness and promote national economic growth; in terms of social progress, artificial intelligence technology is Social governance provides new technologies and ideas. The application of artificial intelligence in social governance is the most direct and effective way to reduce governance costs, improve governance efficiency, and reduce governance interference. In daily life, deep learning, image recognition, and Artificial intelligence technologies such as AI, voice recognition, etc., have been widely used in smart terminals, smart homes, mobile payments, and other fields. In the future, artificial intelligence technology will play a more significant role in education, medical care, travel, and other fields closely related to people's lives. Provide ordinary people with life services with wider coverage, better experience, and better convenience.

IV. Conclusion

To sum up, with the rapid development of Internet technology and the expansion of cyberspace, the distance between countries around the world is getting closer. At the same time, the emergence of various artificial intelligence enriches cyberspace and makes up for the deficiency of real space. However, it also promotes the development of cyberspace structure towards diversification and intelligence, which is more complex. There is no doubt that as the product of social development and technological innovation, the emergence of artificial intelligence technology has directly promoted the progress of global human civilization. This paper presents a student real name data information model based on artificial intelligence technology, which combines cluster analysis and data cloning algorithm, and greatly improves the feasibility, effectiveness and accuracy of student real name data information system.

References

[1] N. Ahmad, U. A. Mokhtar, W. Fariza Paizi Fauzi, Z. A. Othman, Y. Hakim Yeop and S. N. Huda Sheikh Abdullah, "Cyber Security Situational Awareness among Parents," 2018 Cyber Resilience Conference (CRC), Putrajaya, Malaysia, 2018, pp. 1-3.

[2] K. Bhat; V. Sundarraj; S. Sinha; A. Kaul, "IEEE Cyber Security for the Smart Grid," in IEEE Cyber Security for the Smart Grid , vol., no., pp.1-122, 16 Sept. 2013, doi: 10.1109/IEEESTD.2013.6613505.

[3] T. T. Teoh, Y. Y. Nguwi, Y. Elovici, N. M. Cheung and W. L. Ng, "Analyst intuition based Hidden Markov Model on high speed, temporal cyber security big data," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 2017, pp. 2080-2083.

[4] C. Onwubiko and A. Onwubiko, "Cyber KPI for Return on Security Investment," 2019 International Conference on Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), Oxford, UK, 2019, pp. 1-8.

[5] F. Skopik and S. Filip, "Design principles for national cyber security sensor networks: Lessons learned from small-scale demonstrators," 2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), Oxford, UK, 2019, pp. 1-8.

[6] E. Kritzinger and S. von Solms, "Home user security- from thick security-oriented home users to thin security- oriented home users," 2013 Science and Information Conference, London, UK, 2013, pp. 340-345.

[7] M. Frank, M. Leitner and T. Pahi, "Design Considerations for Cyber Security Testbeds: A Case Study on a Cyber Security Testbed for Education," 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech), Orlando, FL, USA, 2017, pp. 38-46.

[8] T. M. Mbelli and B. Dwolatzky, "Cyber Security, a Threat to Cyber Banking in South Africa: An Approach to Network and Application Security," 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), Beijing, China, 2016, pp. 1-6.

[9] K. N. Sevis and E. Seker, "Cyber warfare: terms, issues, laws and controversies," 2016 International Conference On Cyber Security And Protection Of Digital Services (Cyber Security), London, UK, 2016, pp. 1-9.

[10] M. Swann, J. Rose, G. Bendiab, S. Shiaeles and F. Li, "Open Source and Commercial Capture The Flag Cyber Security Learning Platforms - A Case Study," 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Rhodes, Greece, 2021, pp. 198-205.