Eng Int Syst (2022) 6: 423–429

Engineering
Intelligent Systems

Construction of Network Security Job Service Model Based on a Rough Set Data Analysis Algorithm

Jinshan Lin[1,a,*], **Siyu Chen**[2,b], **and Leisong Chen**[3,c]

[1] *College of Information Engineering, Putian University, Putian 351100, Fujian, China*
[2] *School of Management, Putian University, Putian 351100, Fujian, China*
[3] *School of Journalism and Communication, MinnarNormal University, Zhangzhou 363000, Fujian, China*

Since the advent of the Internet, resource sharing and information security have become a pair of irreconcilable contradictions. With the sharing of computer network resources more and more widely, the problem of information security has become increasingly prominent. Therefore, traditional passive protection cannot adapt to the current security situation, and the active network security assessment theory emerges as the times require. This paper reports on the construction of a network security evaluation model based on rough set theory to meet the needs of large-scale network security assessment. The model is used to design and experiment on the network security evaluation system. Experiments show that the model can effectively reduce redundant attributes and can objectively determine the importance of attributes according to the given data. The research results reported in this paper have good practical value and strong operability for the network security management department to carry out continuous network security assessment.

Keywords: Resource Sharing; Information Security; Rough Set; Network Security Evaluation

1. INTRODUCTION

With the rapid development of the information age, the scale of the Internet continues to expand, covering almost all fields of our daily work, life, economy, military, science and education, and even the details of people's clothing, food, housing and transportation are increasingly inseparable from the Internet. However, in the face of the increasingly complex structure and scale of the Internet, as an important basic technology, network security has become an important factor affecting national strategic development and social development. In particular, malicious intruders use a large number of new attack methods to exploit the vulnerabilities of network security, and information systems are facing increasingly serious security risks and threats [1]. Therefore, how to carry out the effective security risk management of information systems has become the focus of research.

It is the premise and foundation of information system security risk management to analyze the security of different stages of the system, conduct an objective and effective security assessment on the information system, and actively discover the potential security risks in the system. It is extremely important to guarantee the security of a network to ensure its normal operation [2]. In recent years, many network security assessment products have appeared on the market, such as intrusion detection systems and vulnerability scanning systems, which play a certain role in strengthening the security of network systems. However, on the one hand, due to the massive amount of security data generated by various security devices, it is difficult to comprehensively and effectively manage and analyse this massive amount of

*Corresponding author: [a]Email:13850260002@139.com, [b]Email: 3451146425@qq.com [c]Email:37939555@qq.com.

information manually, so network security managers cannot ensure the overall security situation of the system; on the other hand, these network security assessment products lack in-depth mining of network security evaluation data CLARIFY WHAT YOU MEAN. Most are limited to the detection and analysis of network security vulnerabilities, so it is difficult to form an overall understanding of the network security situation.

The concepts of attribute reduction and attribute importance in rough set theory are all data-driven, so the evaluation model based on the theory can effectively reduce redundant attributes and determine the importance of the attributes objectively, according to the given data, to improve the accuracy and objectivity of network security assessment. This paper reports on the construction of a network security evaluation model and the development of a network security evaluation system [3]. The model is a security evaluation model that integrates all kinds of network original data, which can objectively analyze the network security situation from multiple angles, and objectively, accurately and comprehensively reflect the network security situation to guide network managers to develop effective security policies. Network security assessment is an essential feature of network security. All network security construction should start with network security assessment. This research applies rough set theory to the construction of a network security evaluation model, which not only solves the problem that some data in network security assessment are not verified before use, it also realizes the mining of the least effective key data from the massive network data, and establishes a practical, effective and comprehensive network security evaluation model to effectively use the data and accurately reflect the purpose of the network state. The network security evaluation model established in this paper has a good practical value and strong operability for the network security management department to carry out continuous network security assessment. It is a powerful means to check the performance of the network system itself and even the network system, so that the network security management department can correctly and comprehensively understand the network security situation. In order to, make reasonable decisions in the investment of network security, the choice of network security measures, the construction of network security guarantee system and other issues, and finally make the network security construction effectively serve the information construction.

2. NETWORK SECURITY ASSESSMENT

With the development of information technology, computer networks have been widely used in many fields. However, the vulnerability of the computer network information system inevitably brings potential security risks to the system. In recent years, threats have confronted computer networks systems due to a variety of unsafe factors existing in every corner of the computer network, which may result in varying degrees of damage and may even paralyze the computer network. How to effectively reduce threats to network security is an important area requiring further study.

2.1 Quantitative Evaluation Method

The quantitative evaluation method uses numerical indicators to quantify risk assessment, which focuses on the value of assets and the quantitative data. The quantitative evaluation method considers two basic elements: the probability of threat events and the possible losses. The result of the cross-product of these two elements is termed ALE (annual loss expectation). In theory, we can determine the risk level of the event by calculating ALE, and then develop a corresponding security policy. For example, one such quantitative evaluation method first evaluates the value V of the information system, decomposes the information system into various components according to the functional units to be more conducive to the pricing of the information system; then, according to the relevant data, the frequency P of the threat occurrence is calculated. Finally, because each risk has different degrees of harm to each asset, it may never cause any harm, or it may have a severe impact. Therefore, it is necessary to calculate the coefficient of the threat impact (μ). According to the above three parameters, the ALE calculation formula is obtained as follows [4–7]:

$$ALE = V \times P \times \mu \tag{1}$$

The problem with this method is that not all data are reliable and accurate. For some types of threats, data are available. For example, the probability of natural disasters in the assessment area can be estimated based on the previous empirical frequency data, and the probability of some system problems can be estimated according to the frequency of system collapse events. However, for other types of threats, there is no frequency data, so it is difficult to determine the accurate impact and probability. In addition, correct security management and control measures can reduce the possibility of threat events, but wrong security management and control measures will increase the possibility of threat events. In both cases, all threat events are interrelated. This makes the quantitative evaluation process very difficult.

The advantage of the quantitative evaluation method is that it can intuitively express the evaluation results with specific values, which makes the results clear at a glance and is more objective than qualitative evaluation methods. The quantitative evaluation method can make the research conclusion more scientific, intuitive and easy to understand. The disadvantage of the quantitative evaluation method is that after the quantification of the evaluation conclusion, although the original complex conclusion is simplified, the conclusion is also fuzzy, and sometimes misunderstanding and misinterpretation may occur and the numerical value is unreliable and inaccurate. At present, the network information system is changing rapidly. It is almost impossible to determine the distribution state function of threats in network security by collecting enough data. Therefore, risk aggregation becomes another problem of quantitative assessment methods. However, it is more difficult to determine the value of the risk associated with each event by aggregating the values of risk together. In a word, the quantitative evaluation method not only improves the accuracy of the evaluation results, it also increases the difficulty of risk aggregation.

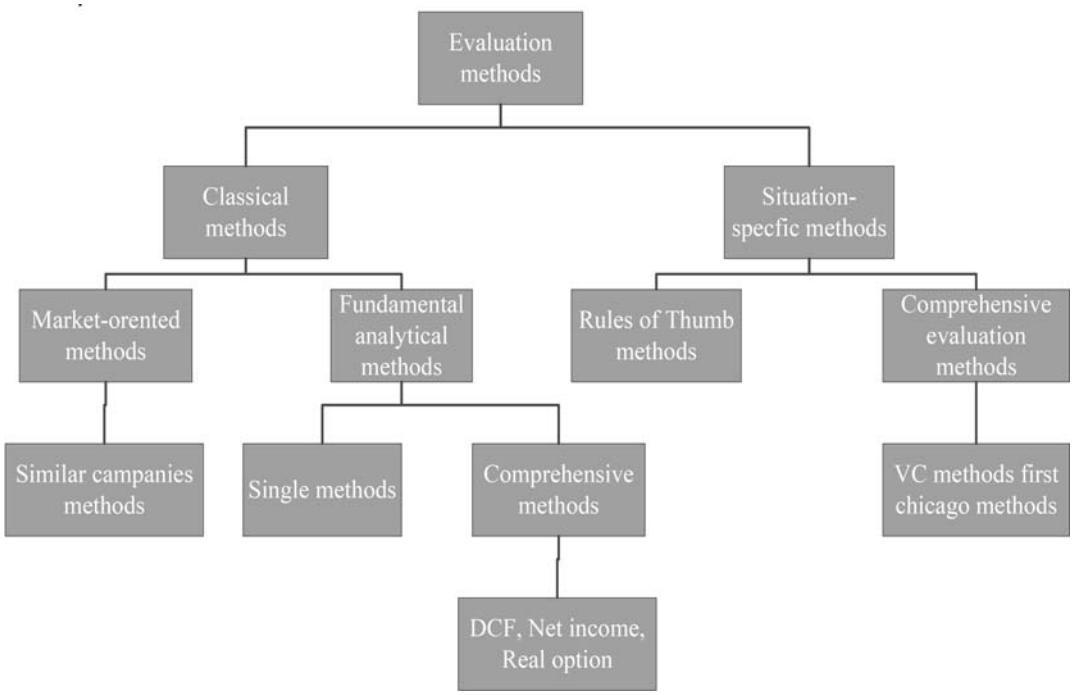

Figure 1 The comprehensive evaluation method.

2.2 Comprehensive Evaluation Method

In reality, especially for a complex network system, a single qualitative or quantitative evaluation method cannot comprehensively or accurately evaluate the security of the network system. Therefore, we need to combine the qualitative and quantitative evaluation methods to suit the situation, which is discussed in this section on the evaluation method. Researchers tend to adopt a comprehensive assessment method combining qualitative and quantitative methods. Conducting a risk assessment of a complex information system is a very difficult process, and various and multi-level factors need to be considered. Some factors can be assessed by quantitative methods, while others cannot or are difficult to quantify, and qualitative assessment methods can be used. However, in the process of the risk assessment of complex information systems, we cannot simply separate qualitative analysis and quantitative analysis, but can only selectively adopt a quantitative method and qualitative method to suit the situation. Only by combining the quantitative method with qualitative method and adopting comprehensive evaluation method can we get scientific and accurate evaluation conclusion. Figure 1 shows the evaluation method.

2.3 Model-Based Assessment Method

A model-based evaluation method is also an effective way to evaluate the security of a complex network system. We can obtain all the possible states and behaviors of the system through the model. The model-based assessment method not only analyzes and identifies the risk factors in the network system, it also finds abnormal or harmful factors in the interaction between the network system and the external environment to qualitatively analyze the weak links and the security threats to the network system. The advantage of the model-based evaluation method is that it is easier to establish the model than to extract rules, and the overall evaluation of the system can be realized by constantly modifying and perfecting the model itself: comprehensively reflecting the security risks in the network system, discovering unknown attack patterns and system vulnerability, etc. The model-based assessment method has gradually become one of the key research directions in the field of risk assessment at home and abroad.

3. ESTABLISHMENT OF NETWORK SECURITY EVALUATION MODEL BASED ON ROUGH SET THEORY

According to the basic viewpoint of rough set theory, "knowledge is a kind of ability to classify objects" [8]. Here, "object" refers to anything we can think of, such as objects, states, concepts, time and space [9–10]. In other words, corresponding to knowledge, there are various classification patterns for the scope of specific things in the concrete or abstract world. Such a scope of specific things is called universe, which is usually a nonempty finite set, and there is no special assumption about the domain and the characteristics of knowledge [11, 12]. Knowledge is the sum of knowledge and experience acquired by human beings in the practice of transforming the objective world. It is abstract, universal and regular. It is also the criterion for people to guide their own behavior. In intelligent decision-making, knowledge is a very important concept, and all decisions depend on knowledge.

In rough set theory, knowledge is defined as the division of the universe, which can be regarded as the ability to classify objects in the universe. The classification ability can be measured by the granularity of knowledge. This is because of the imprecision caused by the granularity of knowledge.

$$M_i = \sum_{t=\tau i}^{\tau i+1-1} lg\left[P(S_t|S_{t-1})P(o_t|S_t)\right] \tag{2}$$

Let u be a nonempty finite set of objects of interest, which is called a universe.

$$M_i = \sum_{t=\tau i}^{\tau i+1-1} lgP(q|o_t) = \sum_{t=\tau i}^{\tau i+1+1} lg\frac{P(o_t|q)P(q)}{\sum\limits_{q_j\in Q} P(o_t|q_j)P(q_j)} \tag{3}$$

where P is the set of all phonemes.

Any subset $x \subseteq U$ of universe u is called a concept or category of universe U. To normalize, we think that an empty set is also a concept called empty concept. Any subset cluster (concept cluster) in universe u is called abstract knowledge about u, or knowledge for short. Each concept (subset) in the universe represents its information granule. In rough set theory, we discuss knowledge which can form a partition or cover on universe U. Usually, in the process of problem solving, we do not deal with a single partition (knowledge) on universe u, but rather a cluster of partitions on universe u, which introduces the concept of a knowledge base.

Rough set theory embeds the knowledge for classification into the set of classical set theory as a part of the set. According to the existing knowledge, we can judge whether an object a belongs to set X. The results can be divided into three cases: (1) object a must belong to set X; (2) object a must not belong to set X; (3) object a may or may not belong to set X. The division of sets is relative rather than absolute, which is closely related to our knowledge about the universe.

Attribute reduction includes attribute reduction and attribute value reduction. Attribute reduction is to delete irrelevant or unimportant attributes under the premise of keeping the classification ability of information system unchanged; the goal of attribute reduction is to find some necessary condition attributes from the conditional attribute set, so that the classification of the decision-making attribute and the shape of all conditional attributes formed according to this part of condition attributes are formed. The classification of decision attributes is consistent. Attribute value reduction further simplifies the decision table on the basis of attribute reduction and obtains a more simplified decision table by removing redundant information from the decision table.

Reduction is related to the consistency of the decision table. When all the decision attributes in the decision table depend on the conditional attribute set, that is, when $C => \mathrm{D}$, the decision table $t =< u, a, C, d >$ is consistent; otherwise, the decision table is inconsistent. According to rough set theory, the consistency of a decision table determines which reduction algorithm to use. The consistency of a decision table can be judged by calculating the dependence between the condition attribute and the decision attribute.

At present, how to get the minimum reduction is a hot issue in the research on rough set theory. Finding the minimum reduction of an information system is a NP complex problem, that is, if the number of objects studied by an information system is m and the number of attributes is n, then the comprehensive reduction of the attribute set needs $(22n \times m)$ basic operations, and the computational complexity increases exponentially with the increase in the attribute table. In the reduction algorithm, the most intuitive is the deletion method, that is to delete the attributes from the data table in turn. The equivalence relationship between the deleted data table and the original data table is compared, and a decision as to whether to delete the attribute or not is determined according to the change of the equivalence relationship. The algorithm can obtain satisfactory results when the number of conditional attributes and records in the information system is small. With the increase of attributes and records, the complexity becomes very high. The heuristic attribute reduction algorithm firstly obtains the initial reduction based on experience, and then adds attributes on the basis of the initial reduction according to the importance of other attributes until the attribute set obtained is the same as the classification ability of the original information system. Then, the attributes in the reduced attribute set are checked one by one, and the attributes that do not change the set's dependence on decision attributes are eliminated.

The criterion of attribute importance is the simplification heuristic factor, which is the core of the algorithm. In rough set theory, knowledge is represented by an information system (attribute value pair table). In general, columns in a table correspond to different attributes, and rows correspond to objects in the universe. The information system is called a decision table when the attributes are further divided into condition attributes and decision attributes. Knowledge reduction judges whether knowledge in information system is necessary for decision-making. In practice, knowledge reduction is the process of deleting redundant information in a decision table until the decision table (approximate) is as small as possible and has the same classification ability as the original decision table. Knowledge reduction includes attribute reduction and attribute value reduction. The reduction of a decision table can also be referred to as the relative reduction of knowledge, and the final result is to reduce the knowledge in the decision table into relatively few decision rules.

In rough set theory, the object in the universe is described by decision table. A decision table is a two-dimensional table, where each row represents an object, and each column represents an attribute of the object. Attributes are divided into two types: conditional attributes and decision attributes. Objects in the universe are divided into decision classes with different decision attributes according to different conditional attributes. For each classification, some conditional attributes are unnecessary, that is, deleting these unnecessary attributes will not affect the original classification effect. Therefore, attribute reduction can be defined as the minimum conditional attribute set that guarantees the classification effect and does not contain unnecessary attributes. A decision table can be reduced in many ways, so there may be many different reduction results. We call the intersection of all these reduction results the core of the decision table. It is obvious that the attributes in the core are important and

affect classification. In other words, the decision table is actually a set of logical rules. Each object in the table contains a classification rule. The two important concepts of reduction and kernel are the quintessence of the rough set method. Rough set theory is the method of solving attribute reduction and kernel. Although the complexity of computational reduction is a typical NP complete problem, it is not necessary to find all the reductions in practice, and it is also not necessary to find the minimum reduction. Sometimes approximate minimum reduction (better reduction) is enough to solve the problem. The best reduction can be found using the heuristic search method, that is, the reduction with the least conditional attributes.

4. IMPLEMENTATION AND EXPERIMENT ANALYSIS OF SECURITY POLICY MODULE

This paper establishes a security policy evaluation model management, and proposes a security policy graded evaluation scheme. The model can be quantified and adjusted in combination with previous experience. Also, a relatively fixed security policy evaluation output template is established to assist the output of the security policy evaluation report of each unit.

4.1 Implementation of the Vulnerability Identification Module

The vulnerability identification module can dynamically generate a security level notification by scanning the vulnerabilities in the network system and combining these with the asset list obtained from the asset management subsystem. The data are shown in Table 1.

4.2 Threat Identification Module

Establish a security hardening plan based on network security experience and business requirements. The threat assessment model management module analyzes all kinds of security incidents and viruses and the degree of harm to the whole network. Combined with the network security experience and the requirements of the business department, the threat assessment model is established, and the threat classification and evaluation quantification scheme is proposed. The security events are quantified according to the harm degree of the events, and the virus situation is counted according to the proportion of virus hosts. The exercise value can be quantified, and the model and quantitative scheme can be adjusted according to the previous evaluation experience. A relatively fixed threat assessment report output template is established to assist the output of the threat assessment report of each unit. The safety event management data table is shown in Table 2.

Virus log data table is shown in Table 3.

4.3 Experiment and Analysis

In this section, the model constructed by rough set theory is used to evaluate the security of information system.. The evaluation process gives new concepts and definition methods. On this basis, a new reduction algorithm is proposed, which obtains a reduction result closer to reality, which makes the evaluation result more objective and optimized. Computer networks are usually set up by large organizations to meet the special needs of the industry. They are usually composed of multi-level backbone networks and user access networks classified according to regional and specific functions. The security risks of the network come from backbone networks at all levels, user access networks and terminal computers in the network. The main tasks of network security risk assessment is to assess the network security risks and can be summarized as follows:

(1) to identify a gap between the security situation of backbone network, user access network and terminal computer and the security objectives required by the existing security policies;

(2) to propose effective measures to revise and improve the security strategy.

The specific assessment elements are as follows:

1) Assets: refers to the backbone network equipment, user access network equipment, terminal computer and other hardware parts of the network. This part of the data can be found through the network topology, and this part of the data can be processed in combination with the device data reported at all levels;

2) Vulnerability: the vulnerability of the asset itself, which can be identified using vulnerability scanning equipment;

3) Threat: the possible factors or events that may cause potential damage to assets. Generally speaking, security incidents and their consequences are the important reasons for analyzing threats. Threats always take advantage of the weakness of assets, namely vulnerabilities, to cause damage to assets. This data can be obtained by analyzing the characteristics of events in duty maintenance;

(4) Security control measures: security measures taken to protect assets from loss, that is, the backbone network, user access network and the terminal's configuration and implementation of various security systems to improve security. This data is obtained through questionnaires and data reported at all levels;

(5) The relationship between the assessment elements: assets have value, and the higher the value of the assets, the greater the risk they will face. If there are exploitable loopholes or vulnerabilities in the assets, they may face the threat of being destroyed. To reduce the occurrence of threats and security risks, a series of security control measures should be implemented.

Table 1 The data for the vulnerability identification module.

Time	Vulnerability identification							
	g	Ʊ	d	SIL	t	Ʊ	s	i
Start time	22	28	35	38	46	51	65	76
End time	27	34	37	45	50	64	75	89

Table 2 The safety event management data.

Name	Data type	Explain
CRSSSBID	Auto number	Security incident management serial number
BJLY	\\ˆrchar (30)	Alarm source
SJMC	Wrchar (30)	Event name
SJXZ	MEMO	Nature of the event
SCSJSJ	TIME	Time of first event
MCSJSJ	TIME	Time of last event
SJYIP	Wrchar (30)	Event source IP address
SJYDW	Wrchar (30)	Unit of event source
SJYOS	Wrchar (30)	Event source operating system
SJYFBDYJ	Wrchar (30)	Install antivirus software on event source host
SJMDIP	Wrchar (30)	Event destination IP address
SJMDDW	Wrchar (30)	Unit of the purpose of the event
SDMDOS	Wrchar (30)	Event stop operating system
SJMDFBDYJ	\ˆrchar (30)	Event destination host installs antivirus software
CLLS	Wrchar (30)	Treatment flow
FKSJ	TIME	Feedback time
FKRY	Wrchar (30)	Feedback staff
FKQK	MEMO	Feedback

Table 3 Virus log data.

Name	Data type	Explain
CRSSSBID	Auto Number	Serial number of virus log management
BJLY	Wchar (30)	Virus killing time
SJMC	Wrchar (30)	Name of infected virus
SJXZ	MEMO	Infected host IP
SCSJSJ	TIME	Infected host operating system

The network security evaluation model based on rough set theory is applied to the large-scale computer network of the company to conduct the network security risk assessment experiment. This paper only discusses the experiment results of the network, which comprises 5 switches, 14 database servers, 2 web servers and 320 office terminals. We collate the security status of the above 341 network assets one month before and after adopting the network security assessment model based on rough set theory, as shown in Table 4. Compare those assets that are considered security threats.

As shown in Table 4, after the network security assessment model based on rough set theory is adopted, the number of key network equipment assets deemed to be in danger of a security threat, such as switches, database servers and web servers will decrease. Of the seven office terminals that are still in danger after the network security assessment based on rough set theory, it is not possible to install anti-virus software or system patches on four office terminals due to hardware issues, and it is not possible to install anti-virus software on three office terminals due to incompatibility issues between the business software and anti-virus software, which results in seven office terminals remaining in a dangerous state. Generally speaking, according to the experiment data, the model greatly improves the security status of the existing network, improves the efficiency of network security management, and has a strong practical significance.

5. CONCLUSION

The network security evaluation model, the selection of evaluation standards, the evaluation implementation process and so on are the research focus of network security evaluation. This paper discusses the network security evaluation system, the network security evaluation process and architecture, and proposes a network security evaluation model based on rough set theory. This paper proposes a network security assessment process which divides network security assessment into four stages and eight processes and establishes a network security assessment model based on rough set theory. In this paper, the network security evaluation model based on rough set theory is applied to a large-scale office network, which greatly improves the security level of the network.

Table 4 Comparison of assets deemed to be in danger.

Asset class	Number	Traditional model		Network security assessment based on rough set theory	
		Number of assets in danger	Percentage of assets in danger	Number of assets in danger	Percentage of assets in danger
Switch	5	1	20*	0	0%
Database server	14		7*	0	0%
Web server	2	0	0≪	0	0%
Office terminal	320	29	9*	7	2%

ACKNOWLEDGMENT

This paper is supported by National Natural Science Foundation of China: Research on Key Technology of Adaptive Bidirectional Relay Communication between Power Line and Wireless Dual Media (No.61601182).

REFERENCES

1. Liu, L. D. (2002) "The Enlightenment of Problem-based Learning on Teaching Reform", Education Research, (2), pp.73–77.
2. Zheng, X. M., & Jiang, Q. Y. (2005) "A Study on Teacher Belief in College English Teaching Reform", Foreign Language Circles, (6), pp.16–22.
3. Zhou, Y. Q. (2003) "The Construction of Excellent Course Materials is an Important Measure of Teaching Reform and Innovation", China Higher Education Research, (1), pp.12–12.
4. Li, Z. Y., Zhu, H., & Liu, Z. J. (2014) "Guiding the Teaching Reform of Higher Engineering Education with the Concept of Achievement Oriented Education", Higher Engineering Education Research, 000(002), pp.29–34.
5. Liu, L. D. (2002) "The Enlightenment of Problem-based Learning on Teaching Reform", Education Research, (2), pp.73–77.
6. Zheng, X. M., & Jiang, Q. Y. (2005) "A Study on Teacher Belief" in College English Teaching Reform", Foreign Language Circles, (6), pp.16–22.
7. Zhong, Q. Q., & Jiang, M. L. (2004) "Value Orientation and Path of Teaching Reform under the Background of New Curriculum", Education Research, 025(008), pp.32–36.
8. Liang, D. F. (2001) "My View on Foreign Language Teaching Reform", Foreign Language Teaching Theory and Practice, (1), pp.8–11.
9. Huang, J. B., & Shao, Y. Z. (1998) "The Way out of College English Teaching Reform", Foreign Language Circles, (04), pp.20–22.
10. Wang, Z. P., & Zhuang, H. H. (2001) "Emphasizing Fitness and Neglecting Competition – Reform and Practice of College Physical Education", Sports Science, 021(001), pp.22–25.
11. Ye, L. (1997) "Let the Classroom Radiate Vitality – on the Deepening of Teaching Reform in Primary and Secondary Schools", Education Research, (09), pp.3–8.
12. Hu, W. Z., & Sun, Y. Z. (2006) "Highlighting Discipline Characteristics and Strengthening Humanistic Education – on Current English Teaching Reform", Foreign Language Teaching and Research, 38(005), pp.243–247.