TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.4, April 2014, pp. 2816 ~ 2825
DOI: http://dx.doi.org/10.11591/telkomnika.v12i4.4288

Network Intrusion Detection System Based on Optimized Fuzzy Rules Algorithm

Liang Lei
QingDao Hotel Management College, QingDao ShanDong 266100, China
email: lei_liang0102@163.com

Abstract

As computer networks and distributed applications more complex, diverse and intelligent, network behavior anomaly detection has gradually become the effective monitoring and system controlling technology. The paper established a network intrusion detection system and to investigate the data rules, sensors and abnormal behavior automatic identification in this system, a kind of algorithm based on fuzzy rules to describe the network abnormal behavior was introduced into this paper. It was used to describe the misclassification invasion rules effectively and then to convert the misclassification invasion rules to the issue of seeking optimal separating hyper plane. Subsequently, the double super ball membership function was introduced into the system to restrict the intrusion features, and to establish intrusion rule set which was used to make optimized description of the intrusion rule set and then complete intrusion detection. The experimental results showed that: in the context of different network attacks, the system can complete a variety of attacks and efficient detection. The detection error was not more than 1% which basically met the requirements of the reliable, high precision, anti-interference ability in automatic network intrusion detection and provided a reference to the future research on network intrusion detection.

Keywords*: computer network, integrated intrusion detection system, fuzzy rule description*

1. Introduction

With the rapid development of computer and network technology, computer networking technology has penetrated into the social, political, cultural, economic, military area and other aspects of people's working life. The impact of it is also growing. Meanwhile, on the other hand, because of the openness and sharing characteristics of computer network, network security issue such as hacking incident appears frequently, which cause a great threat to national security, economic, and social life. Therefore, how to protect the security of the system, develop an appropriate computer network security technology and corresponding measures, becomes the focus of researchers [1-3].

In order to overcome the defects of the current system, it is necessary to establish an intrusion detection system with good adaptability, scalability, flexibility, intelligent, low false positive rate and low false negative rate. Researchers have used a variety of methods to build mathematical models and intrusion detection systems. Fuzzy intrusion detection is one of a modeling approach, which uses fuzzy math and fuzzy data mining to build fuzzy analysis engine to achieve intrusion detection. At present, fuzzy intrusion detection research and development is still at a preliminary stage [4]. The existing intrusion detection systems based on fuzzy rule often use the knowledge of expert in order to prepare detection rules. This artificial rules obviously have great subjectivity and uncertainty, and with the changes in the network environment. These rules do not fulfill the kinds of changes, and thus adaptability is poor [5, 6]. To solve this problem, we propose an optimized fuzzy rule to describe the intrusion detection system. The experimental results show that this is an effective fuzzy intrusion detection attempts. It not only provides a reference for the future in-depth study of intrusion detection technology and provides a theoretical and technical support for establishing information security system.

Received August 10, 2013; Revised October 23, 2013; Accepted November 20, 2013

2. The Detection System

2.1. Composition of the System

Normally the intrusion detection system is composed of eleven components-data source, sensor, behavior, analysis, events, alert, manger, alarm, response, admin and operator. For the host (such as Web server) which is easy to be attacked by hackers in the system, because it can use host-based intrusion detection technology, the installed IDS component can detect the decrypted data to protect the machine from the intrusion [7-9]. Network-based intrusion detection technology can also be applied on the host, so that only detecting the host-related data transmission can prevent hacker attacks, the cost of detection is relatively small in this way. Additionally, this detection method can benefit other application-based intrusion detection and very effective for some hosts in the exchange networks [10-12]

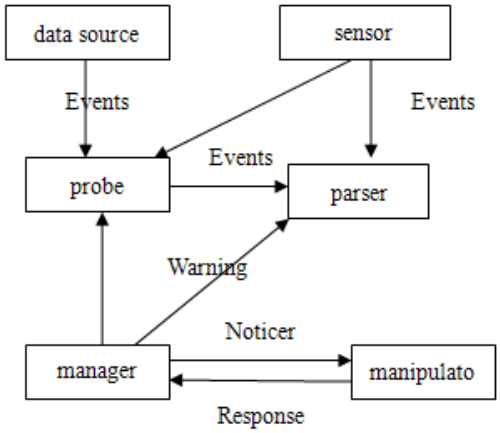

Figure1. Detection System Components

In the network that shared transmission media, there are multiple intrusion detection systems can be set in the key network sections and some hosts need to install HIDS component to ensurethe data will not be processed by NIDS component. The purpose is to avoid duplication of data processing to release the NIDS component workloads.

For the high speed network section, several NIDS in the same network section can be fairly loaded to avoid the “Flooding” denial of service attack (DDOS), and process the captured data packets in section. When there are some faults of the host and the HIDS component can’t work properly, the NIDS in the same section can replace HIDS to capture and detect the data from the host, at the same time, alert the likely intrusion immediately, which can greatly increase the stability of the intrusion detection system. If there were some faults of the NIDS component, the KIDS component in the same network section can be used to detect by changing the range of capture. When a system control component is set in the system, the encrypted one can communicate with IDS component safely. The encryption is for effective management of IDS and avoiding cooperative attacks.

The main functions of the system are:

(1) Manage, control and configure the IDS component;

(2) Collect, analyze and evaluate the detection results from different components in order to make the component can response accurately.

2.2. The System

Detection system work flow chart, with Figure 2 description.

In order to adapt to current complex network environment, we designed a security defense system for multiple attack modes. The operation characteristics and methods of the safe integrated intrusion detection system are as follows.

(1) Embedded operating mode. It can defense the attacks, discard suspicious packages in real time and stop the following data communication.

(2) Availability and reliability. When there are some faults, it can be replaced by the other system detector to maintain the system.

(3) Low latency. The data packages can be processed rapidly and the delay among the link layer, network layer and overall equipment are close.

(4) High performance. The rate required by the practical environment equipment is the same as the data processing capability. When all of the rules are open, fps could meet all of the above requirements.

(5) High intrusion detection accuracy rate. There is no need to restart to apply the characteristic rules rapidly. When some operation in the control center is applied, the corresponding rules are effective in the detectors.

(6) Fine-grained control to stop malicious communication.

(7) Alert process and forensic analysis capabilities. When the sensors alert, the surveillance center can provide real-time and previous records to find whether there are some correlations and determine how to response.

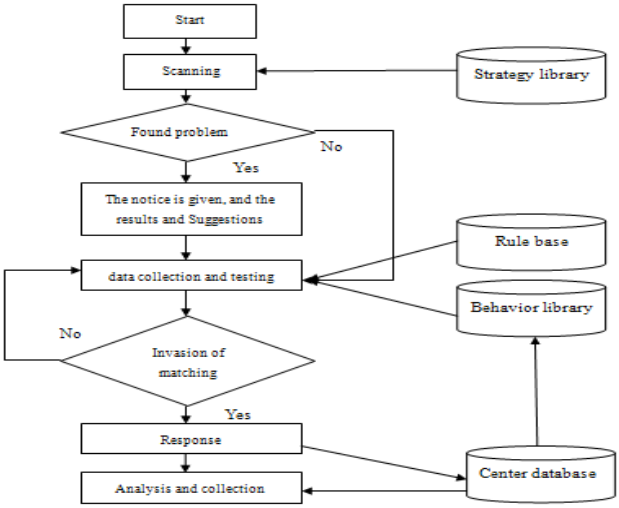

Figure 2. Detection System Work Flowchart

Figure 3 is the scheme of the network intrusion detection system. The operation modules are classified by system logic based on the operation stage and tasks. There are two stages for the detection system-training and testing. The training stage is to analyze the training dataand store the mode rules into SQL. The testing stage is to analyze the captured data with SQL and further process the analyzed results by the response module. All of the above tasks are finished by central control module.

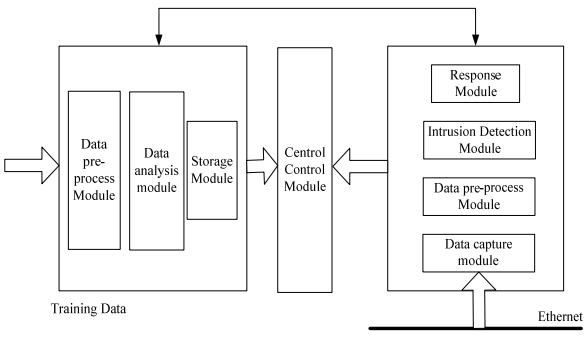

Figure 3. Scheme of the Network Intrusion Detection System

The system control center is the core of the whole model. The control logic of the whole system is generated by the module. The control center distributes and dispatches the separated function logic to increase the operation efficiency, and benefit further update and maintenance. The control central defines multiple functions in the system, such as normal function in the control interface, interior function in the control module, and the operation functions among modules. The control central pre-defines several classes- capture class, pre-process class, analysis class, and response class. The above classes can be dispatched and managed by visible management interface. The main functions are: ① test unit management; ② receive alert information and display; ③ log data management; ④ user management; ⑤ rules definition.

3. Fuzzy Invasion Characteristics Rules

3.1. Fuzzy Rules of Training Set

Assuming the training set of intrusion detection system is $S=\{(x_1,y_1,u(x_1)),...,(x_l,y_l,u(x_l))\}$, in which $x_j \in R^n$, $u(x_j) \in \{-1,1\}$, $\sigma \le u(x_j) \le 1$, σ is real number which is greater than 0. $u(x_j)$ is the training point, the output is $(x_j, y_j, u(x_j))$ $y_j = 1$ (positive) or $y_j = -1$ (negative), the fuzzy membership is $(j=1,...,l)$. The fuzzy membership $u(x_j)$ is the degree of training point $(x_j,y_j,u(x_j))$ belonging to a certain kind, the parameter ξ_j is the measurement of the misclassification degree, so $u(x_j)\xi_j$ has become the measurement for a measure of variable wrong indexing of different importance. For linear problems, finding the optimal separating hyperplane is changed into solving a quadratic programming problem as follows:

$$\left.\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{j=1}^{l} u(x_j)\xi_j \\ &s.t. \quad y_j((w \cdot x_j) + b) + \xi_j \ge 1 \\ &\xi_j \ge 0, \quad j = 1,2,\ldots,l \end{aligned}\right\} \tag{1}$$

In which $C>0$ is punishment parameters, $\xi = (\xi_1,...,\xi_j)^T$, $u(x_j)$ is the degree of training point $(x_j,y_j,u(x_j))$ belonging to a certain class. Solving the dual programming of quadratic programming (1). According to the dual definition of $Wolfe$, getting minimum value of $Lagrange$ function with w, b, ξ_j as follows.

$$\begin{aligned} &\frac{\partial L(w,b,\xi,\alpha,\beta)}{\partial w} = w - \sum_{j=1}^{l}\alpha_j y_j x_j = 0 \\ &\frac{\partial L(w,b,\xi,\alpha,\beta)}{\partial b} = -\sum_{j=1}^{l}\alpha_j y_j = 0 \\ &\frac{\partial L(w,b,\xi,\alpha,\beta)}{\partial \xi_j} = u(x_j)C - \alpha_j - \beta_j = 0 \end{aligned} \tag{2}$$

Bring Equation (2), seeking the maximum of α to find the dual programming of great Quadratic.

$$\left.\begin{aligned} &\max_{\alpha} \sum_{j=1}^{l}\alpha_j - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) \\ &s.t. \quad \sum_{j=1}^{l} y_j \alpha_j = 0 \\ &\quad 0 \le \alpha_j \le u(x_j) C, \qquad j = 1, 2, ..., l \end{aligned}\right\} \tag{3}$$

The seeking of optimal hyperplane problem is transformed into solving quadratic programming (1) dual Planning (3)-ended questions. Planning (3) is a convex quadratic programming solution, the optimal solution is $\alpha^* = (\alpha_1^*, ..., \alpha_l^*)^T$, so the fuzzy optimal classification function is as follows.

$$f(x) = \operatorname{sgn}\{(w^* \cdot x) + b^*\}, x \in R^n \tag{4}$$

In which, $w^* = \sum_{j=1}^{l} \alpha_j^* y_j x_j$, $b^* = y_i - \sum_{j=1}^{l} y_j \alpha_j (x_j \cdot x_i)$, $i \in \{i \,|\, 0 < \alpha_i^* < u(x_i) C\}$.

In $\alpha^* = (\alpha_1^*, ..., \alpha_l^*)^T$, only part of $\alpha_j^* > 0$ is OK, the input x_j of the corresponding training point is support vector. There are usually two general vectors set which are supported. One is the support vector which is corresponded to $0 < \alpha_j^* < u(x_j) C$, the distribution of the support vector is at the edge of hyperplane the other one is corresponded to $\alpha_j^* > u(x_j) C$, this support vector is misclassified samples. The biggest difference between fuzzy support vector machine and traditional support vector machine is the existence of $u(x_j)$, in fuzzy support vector machine, the corresponding support vector of α_j^* is different from α_j^* in traditional vector machine.

For nonlinear problem, the function of $k(x_i, x_j)$ is used, the classification can be expressed in quadratic programming as follows.

$$\left.\begin{aligned} &\min_{\alpha} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{j=1}^{l}\alpha_j \\ &s.t. \quad \sum_{j=1}^{l} y_j \alpha_j = 0 \\ &\quad 0 \le \alpha_j \le u(x_j) C, \qquad j = 1, 2, ..., l \end{aligned}\right\} \tag{5}$$

Planning (5) is a convex quadratic programming. The optimal solution is $\alpha^* = (\alpha_1^*, ..., \alpha_l^*)^T$, so the fuzzy optimal classification function is as follows.

$$f(x) = \operatorname{sgn}\{\sum_{j=1}^{l} \alpha_j^* y_j K(x, x_j) + b^*\}, x \in R^n \tag{6}$$

$b^* = y_i - \sum_{j=1}^{l} y_j \alpha_j K(x_j, x_i)$, $i \in \{i \mid 0 < \alpha_i^* < u(x_i)C\}$.

Where,

3.2. Double Hypersphere Membership Function

A double hypersphere membership function is introduced into the system. The relationship between the sample and the class center, and the relationship between each sample in class are considered sufficiently when determining the degree of membership. In addition, the sample membership is seen as a non-linear relationship with the distance between the sample and the center of the class. The traditional SVM makes an appropriate dividing of valid samples, noises and outliers, most of the samples located on the one side of classification surface are valid samples while the other side are noises and outliers. The center of sample data x_0 is regarded as the center of sphere, building a cutting ball with the radius R which is the distance between classification surface H and the centre point, denoted as sphere A. Then x_0 is taken as sphere center to build a new sphere with radius 2R, which is sphere B. The samples in sphere A are taken as valid samples, and are given large membership; the samples located outside of the sphere B are regarded as noises and outliers, the membership degree is zero, then, the samples between the two spheresare given a smaller degree of membership to indicate that it is the degree of valid samples. After SVM classifier, according to the classification function, $\{x_j, j = 1,2,3\cdots,n\}$ can be obtained, the center vector is $x_0 = \frac{1}{n}\sum_{j=1}^{n} x_j$, linear discriminant function is $g(x) = w\,x + b$, the distance from the arbitrary point x to the classificationsurface can be expressed as $g(x)/\|w\|$, radius is $R = g(x_0)/\|w\|$.

The S-shaped function is combined to define the fuzzy membership $u(x_j)$ as:

$$u(x_j) = \begin{cases} 1 - \dfrac{d(x_j)^2}{2R^2}, (d(x_j) \le R) \\ \dfrac{[d(x_j) - 2R]^2}{2R^2}, (R < d(x_j) \le 2R) \\ 0, (d(x_j) > 2R) \end{cases} \quad (7)$$

$d(x_j) = \|x_j - x_0\|$ is the distance between x_j and x_0.

3.3. Algorithm Description

Fuzzy Support Vector Machine (FSVM) algorithm makes use of the above principles and formulas to design intrusion detection classifier , input the training data and testing data, output the type of detection data (normal, intrusion or abnormal). The algorithm is described as follows.

1) Train traditional SVM classifier, obtain the initial support vector, and use it to constitute a decision classification surface $w\ x + b = 0$.

2) Calculate the vector of Computing Center x_0.

3) Calculate radius R and 2R according to the decision classification surface and x_0.

4) Calculate membership function $u(x_j)$ according to equation (10).

5) Get fuzzy training set $\{(x_1, y_1, u(x_1)), (x_2, y_2, u(x_2)), ..., (x_l, y_l, u(x_l))\}$.

6) Train fuzzy training points, build the optimal structure of the classification function, and obtain fuzzy support vector machine classifier

4. Optimal Detection Process of the Rules

Fuzzy rules can be obtained by such above methods, but the detection speed of the system is limited by the number of rules, in order to improve the processing speed, the rule set need to be optimized. During the optimization process, the rule set must be constructed and selected.

4.1. Construction of the Rule Set

It is necessary for the rule set detection method to structure the rule set by using the rule optimizer. There are two requirements for the rule optimizer. (1) The rule optimizer must be set to achieve the purpose of constructing smallest and most efficient rule set. (2) Discrete rule set should be structured. Thus, each data packet only needs to search one rule set.

In the initialization process, the rule set is structured with the most independent optimizer Snort rule parameters by the rule optimizer. The different parameters for each type of transmission protocol are independent, so the selected rule parameters for the type of transmission protocol is different. For example, the TCP rule set can be distinguished from each other with the source and destination ports, and the ICMP rule setcan be distinguished from each other according to the rules of ICMP type. A subset is structured with the independent parameters, whichallows the detection engine of multiple rules can detect the smaller rule set. More importantly, it allows the data passing by the corresponding subset of the rules according to the characteristic of the data packets.

4.2. Choosing the Rule Set

When the snort is running, a rule set for each data packet is selected by the rule optimizer. Some of the parameters and rule set are selected depending on the matching results of received packets lumped parameter. Thus, only those rules which match the packet of the rules are select. Since then multiple rules search engines can detect the content by the detection method based on the rule set of the rule set testing methods. For some abnormal packets, maybe two sets of the rules will be selected, this is a condition called “independent conflict”.

4.3. Actual Application of the Rule Optimizer

According to the independent parameters of transmission protocol the rules are divided into definable rules and small rule sets by the rule optimizer to improve the speed of snort detection. By analyzing, the source port and destination port can be used as independent parameters of the TCP/UDP packet. The ICMP types can be used as an independent parameter of the ICMP packets.

Set of rules with the optimized overall structure is constructed, which is more detailed. When a packet is got, first of all, determine whether the IP protocol field is exist, if not, treat it with common IP rules. If yes, judge which the rule is, TCP/UDP, ICMP or others. If it is TCP/UDP rule, it should be processed according to the appropriate rules set of independent parameters, at last, according to the detection result to judge whether it is malicious packets or the invasion.

5. Simulation

5.1. Data Source and Test Environment

The experimental data is from standard data set KDD 1999, there are four kinds of attacks in the data set. They are scanning and probing (Probe). denial of service attacks (DoS), unauthorized remote access (R2L) and the local super user illegal access (R2R), the other data are normal, the 4 types of intrusion data set are selected randomly from the data set as shown in Table 1.

Table 1. Sample Data

intrusion Type	training		test	
	ormal	ntrusion	ormal	ntrusion
Probe	731	711	576	970
DoS	150	904	640	287
U2R	84	3	57	80
R2L	72	89	24	18

Figure 4 is a intrusion detection environment with Gigabit speed, all of them are PC computers with Windows XP operating system, the processor is P4 dual-core 1.8G, the DDR memory is 1G, the hard disk is 160GB, and the Ethernet card is 1Gigabit.

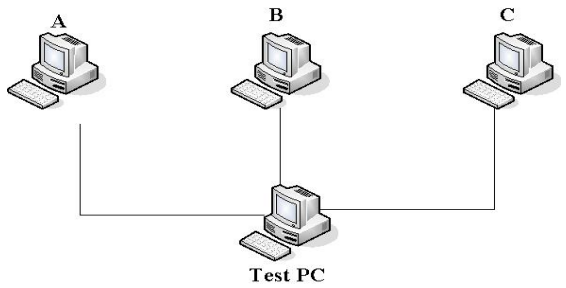

Figure 4. Test Environment

Test card will send Ethernet Frame directly to the user level, without any treatment, in the case of different IP packet size, packet capture performance, as shown in Table 2

Table 2. IP Package Capture

IP package Size byes)	Speed of sending package(pps)	Flow of sending package(bps)	Receiving package speed (pps)	Flow of receiving package(bps)	Receive rate
64	585968	200M	585968	200M	100%
	645875	280M	645875	280M	100%
256	786550	350M	663061	295M	84%
	244141	500M	244141	500M	100%
	400000	600M	324000	486M	81%
1024	97656	800M	97656	800M	100%
	110000	900M	100100	819M	91%

Seen from the data capture package technology with Winpcap, the packet capture platform with Winpcap has better performance than the traditional packet capture platform. It can adapt the large flow of network environment for the high-speed packet reception rate.

5.2. Comparative Models and Evaluation of Model Performance

In order to make the network intrusion detection system more convincing, all eigenvalues and support vector machine model (SVM) are used, the particle swarm individually is selected and the support vector machine parameter model (PSO-SVM) is taken as a comparison model to evaluate the performance with detection rate and run time.

5.3. Test of Feature Selection

The compared result between PSO-SVM algorithm and algorithm in this paper used to select feature of the network intrusion detection is shown in Table 3.

Table 3. Different Models Feature Selection Results

Type of intrusion	Num before choose	PSO-SVM	Algorithm in this paper
Probe	41	17	15
DoS	41	18	17
U2R	41	12	10
R2L	41	14	11

Seen from Table 3 after feature selection algorithm, the number of feature processed with algorithm in this paper is less than that of PSO-SVM, and the number of characteristics is less than the number of pre-select feature, so it is necessary to select the number of the network intrusion features, with which, the number of input variables for the support vector machine can be reduce greatly, and the learning speed of network intrusion detection can be accelerated.

5.4. Comparing of the Detection Result

The selected characteristics are input to the support vector machine to learn with optimal support vector machine modeling, the test sample is checked with the optimal detection model, and the detection result is shown in Table 4. Seen from the comparison result of the table, the model with feature selection has a higher detection rate than model with all the original features, and the detection rate of the proposed algorithm is higher than that of PSO-SVM, the comparison result shows that the combined model of characteristics of selection and support vector machine parameters can take the advantage of each algorithm and dig the network status information.

Table 4. Detection Rate Comparison of Different Models

type	SVM(%)	PSO-SVM(%)	Algorithm in this paper(%)
Probe	99.51	98.31	99.56
DoS	98.38	97.60	99.51
U2R	98.14	98.16	99.71
R2L	98.00	97.64	99.17

6. Conclusion

A characteristics description algorithm of abnormal network behavior detection based on fuzzy representation rules was proposed in this paper. The detail of the misclassification of invasion rules was described and converted into an issue of seeking the optimal separating hyperplane. Furthermore, the double super ball membership degree function was introduced to the system to restrict the intrusion characteristics and the set of invasion rules was established. At last, the set of rules were optimized to complete intrusion detection. The results showed that the designation in this paper was effective, feasible. It not only provided a reference for the future in-depth study of intrusion detection technology, but also provided a theoretical and technical support to establish a security network system.

References

[1] Mohammad Saniee Abadeh, Jafar Habibi, Zeynab Barzegar, Muna Sergi. A parallel genetic local search algorithm for intrusion detection in computer networks*. Engineering Applications of Artificial Intelligence*. 2007; 20(8): 1058-1069.
[2] Roberto Perdisci, Giorgio Giacinto, Fabio Roli. Alarm clustering for intrusion detection systems in computer networks. *Engineering Applications of Artificial Intelligence*, 2006; 19(4): 429-438.
[3] GUO Rong-yan, HU Xue-hui. Study aboutLicense Plate Recognition Based on Back Propagation Neural Network. *Computer Simulation*. 2010; 27(9): 299-301.
[4] Ivan Goethals, Kristiaan Pelckmans, Johan AK Suykens, Bart De Moor. Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*. 2005; 41(7): 1263-1272.
[5] Youping Zhao, Lizdabel Morales-Tirado, Cognitive Radio. Forging ahead from Concept Testbed to Large-Scale Deployment. *Journal of Communications*. 2012; 7(7): 514-523.

[6] Xue Hua, Li Xue-ying, Chen Yu. Research on the intrusion detection and its implementation through Data streaming. *Computer Applications.* 2004; 4(1): 112-114.
[7] Zhu Hong. Fault Diagnosis for Analog Circuits Based on D-S Evidence Theory and PSO Neural Network. *Computer Measurement & Control*. 2013; 21(4): 868-870.
[8] Xiaojian Wu, AL Narasimha. Reddy A Novel Approach to Manage A Hybrid Storage System. *Journal of Communications*. 2012; 7(7): 473-483.
[9] ForrestS, Hofmeyr SA, Somayjia. Computer Immunology. *Communications of the ACM*. 1997; 40(10): 88-96.
[10] Sharief MA Oteafy, Hossam S Hassanein. Resource Re-use in Wireless Sensor Networks: Realizing a Synergetic Internet of Things. *Journal of Communications*. 2012; 7(7): 484-493.
[11] Hammad M, Conor RA. Less Destructive Context-awa-re Crossover Operator for GP. *Berlin/Heidelberg: Springer-Verlag*. 2006; 45(5): 36-48.
[12] Saeid Asgari Taghanaki, Behzad Zamani Dehkordi, Ahmad Hatam, Behzad Bahraminejad. Synthetic Feature Transformation with RBF neural network to improve the Intrusion Detection System Accuracy and Decrease Computational Costs. *International Journal of Electrical and Computer Engineering*. 2012; 1(1): 28-36.