**TELKOMNIKA Indonesian Journal of Electrical Engineering**
Vol.12, No.3, March 2014, pp. 1954 ~ 1959
DOI: http://dx.doi.org/10.11591/telkomnika.v12i3.4400


# Efficient Computer Intrusion Detection Method Based on Artificial Bee Colony Optimized Kernel Extreme Learning Machine

**Zhigang Zhang**
School of Computer Science and Technology, Hubei Polytechnic University
North Guilin Road No 16, Huangshi 435003, China, telp/fax: +86-714-6353390
e-mail: zhigangzhang_hpu@163.com

***Abstract***

*With continuous development of computer networks, network attacks threat the information security of people's daily life. For the protection against network intrusion behaviors, it is imperative to search efficient measurements to maintaining network security. Literature review indicates that taking the advantages of neural network, the network intrusion can be efficiently detected and the kernel extreme learning machine (KELM) can provide quick and accurate intrusion detection ability. The only parameter need be determined in KELM is the neuron number of hidden layer. Suitable neuron number will accelerate the training procedure. However, little work has been done to address the optimization of KELM. To address this issue, this paper proposed an effective method that uses the artificial bee colony (ABC) to optimize the KELM. With proper hidden layer neuron number, KELM could enhance the accuracy and speed of the intrusion detection. To verify the proposed method, experimental tests have been implemented in this work. The test result demonstrates that the proposed ABC-KELM can detect the network intrusion efficiently and its performance is superior to unoptimized KELM method.*

***Keywords****: computer science, network intrusion, intelligent detection, artificial neural network*



## 1. Introduction

Computer and Internet technology have changed human's life. Along with the development of the Internet, network security issues are increasingly becoming prominent [1]. The computers are now suffering from hackers, viruses, malware and other misconduct [2]. As a last line of security defense, the intrusion detection system can be used to detect various forms of intrusion behaviors. In the protection against the threat of network intrusion, the intrusion detection system is particularly important and necessary for scientific construction of network security [3].

The Internet has provided great convenience for computer users in information sharing; due to the openness of the architecture of Internet, a user's own information and data, including confidential information and data are exposed in the presence of external network users. Hence, the network must have a strong enough security to protect users' privacy in all aspects of different kinds of threats and vulnerabilities [4]. In this way we can ensure network information confidentiality, integrity, and availability. At present, the commonly used network security technologies mainly include firewall, encryption technology, authentication and digital signature technology, anti-virus technology, access control, etc. However, the type of network attack emerges in endlessly and the intrusion means constantly update. It makes the defense system hard to detect many attacks [4]. As a result, more powerful active strategies and solutions will be enhanced to maintain the network security.

In order to enhance the intrusion detection performance, some useful measurement techniques have been proposed. These techniques include the probability statistical model based on user behavior, intrusion detection expert system, neural network based intrusion detection technology, intrusion detection model based reasoning technology, intrusion detection technology based on state transition analysis, intrusion detection technology based on the immune system and so forth [5]. These techniques have been proven to be very useful in the intrusion detection in many applications. However, they still have some problems.

*Received August 17, 2013; Revised September 16, 2013; Accepted October 8, 2013*

(1) Intrusion detection system is still in development and is far from mature. At present, the vast majority of commercial intrusion detection systems are of a similar principle and virus detection and the self-adaptive intrusion detection system is far from mature. Intrusion detection technology remains to need breakthrough in theory.

(2) False alarm rate is high, seriously interfering with the detection results.

(3) Incident response and recovery mechanism is imperfect. This part is very necessary for the intrusion detection system, but at present, is almost ignored.

(4) The collaborativation of different technologies is insufficient in intrusion detection system.

(5) Lack of further explanation and analysis tools for the test results.

The detection model based on neural network hopes to be able to effectively maintain the network security to build a harmonious network environment. The most popular artificial neural networks (ANNs) are BP NN and RBF NN. They are powerful to find hidden information from random data; however, they suffer from local minima and slow convergence speed [6]. The KELM has no such shortcomings and is very efficient in pattern recognition [7]. The KELM is a kind of single-hidden layer feed-forward network (SLFN) and it has only one parameter in the SLFN structure, i.e. the neuron number of the hidden layer [7, 8]. Zong and Huang [9] adopted the ELM into the face recognition. They [7] also employed the KELM in the face recognition and found that the ELM and KELM are more efficient than the BP NN and support vector machine (SVM). However, a parameter optimization mechanism of the KELM or ELM has not developed in their work. Usually, bad value of the neuron number may influence the performance of the KELM [10]. Hence, it is wise to test the optimization of the KELM and its application in the network intrusion detection.

To address the mentioned issue, for the first time, a new method based on artificial bee colony (ABC)-KELM is proposed for the network intrusion detection in this work. The ABC is used to optimize the hidden neuron number of the KELM. Experimental tests have been carried out to evaluate the performance of the ABC-KELM method. The intrusion detection performance of the proposed intrusion detection method is compared with KELM, ELM, BP NN and RBF NN in terms of both intrusion detection accuracy and training speed.

## 2. Research Method

This work will introduce the ABC-KELM method for the intrusion detection. The theories about ABC and KELM are briefly described as follows.

### 2.1. The Artificial Bee Colony (ABC)

Artificial bee colony, proposed by Karaboga [11], is a kind of bionic algorithm inspired by the behaviour of honey bee swarm. Artificial bee colony consists of three types of bees: employed bees, onlookers and scouts. Half of the swarm is employed bees, and the other half is onlookers. Their Numbers are equal to the number of food source, i.e. the solution of the optimization. Every employed bee has a correspondence relationship to a food source. The employed bee quantity to a food source represents the quality of the solution. If a food source is abandoned by all the employed bees and onlookers, the employed bees corresponding to this food source become scout automatically.

ABC firstly generates the population containing *N* food sources. Each solution is $x_i = [x_1 \quad x_2 \quad \cdots \quad \mathbf{s}_N] \in R^D$. Then the employed bees and onlookers circularly search all the food sources. In the searching, the employed bees will search the neighborhoods of their corresponding food sources and then select the better ones according to greed strategy. After one round of neighborhood searching, the new correspondence relationship of each employed bee will transmit to the onlookers. The onlookers will select a food source using probability $p_i$ according to the quality of the food sources to do the same neighborhood searching. Then the best solution will be chosen among the searching results.

The selection probability $p_i$ can be expressed as [11]:

$$p_i = f_i / \sum_{n=1}^{N} f_n , \tag{1}$$

Where $f_i$ is the quality of the *i*th food source.

The neighborhood searching is given by:

$$\lambda_{ij} = x_{ij} + \xi_{ij} \ (i \neq j), \tag{2}$$

Where $i \in [1, 2, ..., N]$, $j \in [1, 2, ..., D]$ and $\xi$ is a random constant.

If an optimum solution is not derived after *M* searching for a food source, the corresponding employed bee will transform into scout and the food source will be initialized.

### 2.2. The Kernel Extreme Learning Machine (KELM)

Given samples $\{(x_i, t_i) : i = 1, 2, ..., N; x_i \in R^p, t_i \in R^q\}$, where *x* is the feature vector and *t* is the class label vector, the below SLFN is used to identify the sample [8].

$$\sum_{i=1}^{m} \beta_i \, g(\alpha_i^T x_j - b_i) = o_i, j=1, 2,..., N. \tag{3}$$

Where, *m* is the number of hidden neuron; $o_i$ is the output of *j*th sample; $g(\cdot)$ is the activation function; $b_i$ is the threshold of the *i*th hidden neuron; $\alpha_i$ and $\beta_i$ are the input and output weight vectors, respectively. If the output $o$ can approximate *t*, we derive:

$$\sum_{i=1}^{m} \beta_i \, g(\alpha_i^T x_j - b_i) = o_j = t_j, j=1, 2 ,..., N. \tag{4}$$

(4) can be written compactly as:

$$\mathbf{G\beta} = \mathbf{T}, \tag{5}$$

Where,

$$\mathbf{G} = \begin{bmatrix} g(\partial_1^T x_1 - b_1) & \cdots & g(\partial_m^T x_1 - b_m) \\ \vdots & \cdots & \vdots \\ g(\partial_1^T x_N - b_1) & \cdots & g(\partial_m^T x_N - b_m) \end{bmatrix},$$

$\boldsymbol{\beta} = [\beta_1, \cdots, \beta_m]^T$ and $\mathbf{T} = [t_1, \cdots, t_N]^T$.

To solve (5), the ELM adopts a least squares error to get solution $\hat{\boldsymbol{\beta}}$:

$$\hat{\boldsymbol{\beta}} = \mathbf{G}^{\dagger}\mathbf{T}, \tag{6}$$

Where, $\mathbf{G}^{\dagger}$ is the Moore-Penrose generalized inverse of $\mathbf{G}$. Function $g(\cdot)$ is usually unknown, we can incorporate kernel functions in $g(\cdot)$. This is the so called KEML. The kernel matrix $\mathbf{K} = \left[K(x; x_1) \ \cdots \ K(x; x_N)\right]^T$ ($K(\cdot)$ is the kernel function) is introduced into (5) and (6) to estimate the output of the KELM:

$$\mathbf{o} = \mathbf{KT} \tag{7}$$

Herein, the Gaussian kernel function (RBF) is adopted.

$$K(x_1; x_2) = \exp\left(\frac{-\|x_1 - x_2\|^2}{2\sigma}\right), \tag{8}$$

Where, $\sigma$ is the width of RBF.

The number of hidden neuron *m* needs to be optimized for KELM. To do so, the ABC is used to optimize *m* in the training processing of the KELM.

### 2.3. The New Method Based on ABC-KELM

The proposed network intrusion detection method can be summarized as follows:

Step 1: Format the intrusion data into standard form.

Step 2: Fuse the data using principal component analysis (PCA) to obtain the feature vector.

Step 3: Train the KELM using the feature vectors, and optimize the hidden neuron number using ABC.

Step 4: Test the performance of the ABC-KELM detection model. A workflow block of the proposed ABC-KELM intrusion detection method is give in Figure 1.

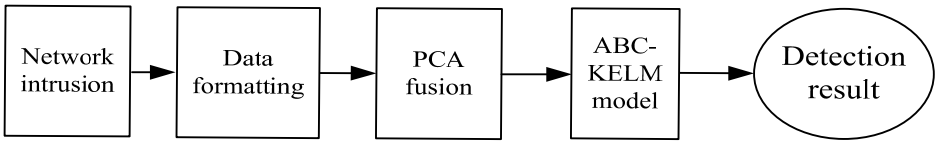


Figure 1. The Workflow Block of the Proposed ABC-KELM Intrusion Detection Method

## 3. Experiments

An experimental platform has been established in the presented work to evaluate the proposed ABC-KELM intrusion detection method. Figure 2 shows the experiment platform, which consists of the server, link connection equipments and host computers. The established experimental platform can be used to test typical network intrusions, including the User to Root (U2R), Remote to Local (R2L), Denial of Service (DoS) and Probe or Scan (PoS) and so forth. In the experiments, we have adopted the U2R, R2L, DoS and PoS to verify the proposed detection method. We have recorded the attacking features of the bytes issued from source to destination, the bytes from destination to source, duration, teardrop, neptune, etc. After the experiments, we have collected 3,000 samples for each intrusion and the total samples are 12,000. Half of the samples is used for training of the KELM, and the rest half is used for testing.

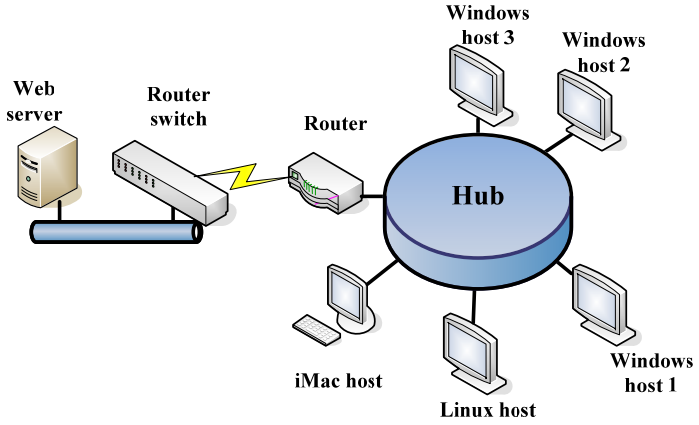


Figure 2. The Principle of the Experiment Tests

The PCA is firstly used to fuse the features into 2, 3, and 8 principal components, respectively. Then the new components are compared in the intrusion detection using the ABC-KELM model. Moreover, the detection performance of the ABC-KELM model is compared with some rivals, including the ABC-ELM model, KELM model, BP NN model, and RBF NN model.

Table 1 lists the PCA analysis results. It can be seen in the table that the first 2 principal components contain 93.32% information of the whole feature space, the first 3 contain 96.43% and 99.68% for the first 8 principal components. These results indicate that the PCA can fuse effecient principal components to present the distinct characteristics of the feature space. Hence, these extracted principal components can be used as inputs into the KELM for the intrusion detection.

Table 1. The PCA Analysis Results for the Feature Fusion

| Principal components | Cumulative variance contribution |
|---|---|
| First 2 components | 93.32% |
| First 3 components | 96.43% |
| First 8 components | 99.68% |

Figure 3 shows the visualization of the first 2 principal components extracted from the 12,000 samples, and Figure 4 shows the visualization of the first 3 principal components. One can be noticed from Figure 3 and Figure 4 that after the PCA processing the distinguished information of different intrusion types can be extracted to identify the intrusion types. It is also can be seen in the two figures that the performance of the PCA processing is not good enough because of overlaps of different intrusion types. It is therefore crucial to employ the KELM to precisely recognize the network intrusions.

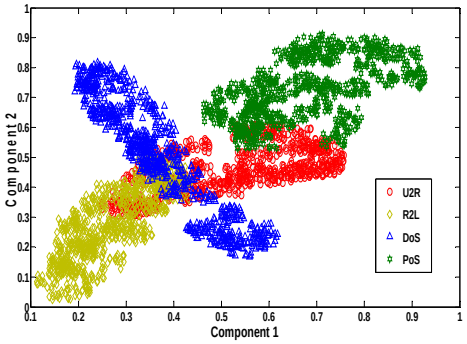


Figure 3. Visualization of the First 2 Principal Components

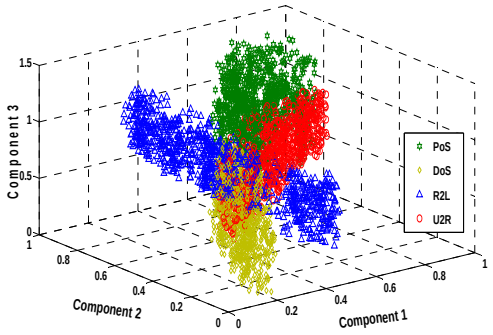


Figure 4. Visualization of the First 2 Principal Components

After the PCA analysis, the KELM is used to recognize the intrusions. Firstly, 6, 000 samples was used to train the KELM, and the reminders were used to test. In the training, ABC was adopted to optimize the parameter of KELM. The intrusion detection results are shown in Table 1. The proposed ABC-KELM method was compared with the KELM, ELM, BP NN and RBF NN. Table 2 shows the comparison results using 2 principal components.

It can be seen in Table 1 that by the PCA analysis, the intrusion detection correction is increased at least by 3.4% and the false positive rate is decreased at least by 4.4%. It also can be seen in the table that the intrusion detection correction will not increase with the increase of the principal components. Hence, the analysis results suggest that the PCA analysis can efficiently enhance the intrusion detection rate and the intrusion detection performance of the proposed ABC-KELM model is acceptable.

In Table 2, one can be noticed that the both the intrusion detection accuracy and training speed of the proposed ABC-KELM method is among the best. The ABC optimization processing can significantly improve the intrusion detection accuracy. Owning to the simply neuron network structure, the training speeds of the KELM and ELM based models are equal but better than that of the BP NN and RBF NN. As a result, the comparison results prove that the proposed ABC-KELM can detect the network intrusion efficiently and its performance is superior to its rivals.

Table 1. The Intrusion Detection Result of the ABC-KELM Model

| Principal components | PCA-ABC-KELM | |
|---|---|---|
| | Detection rate (%) | False positive rate (%) |
| 2 | 89.7 | 7.3 |
| 3 | 90.3 | 6.7 |
| 8 | 90.3 | 7.3 |
| Original feature space | 86.3% | 11.7% |

Table 2. The Comparison between the ABC-KELM, KELM, ELM, BP NN and RBF NN

| Method | Recognition rate | Training time |
|---|---|---|
| PCA-BP NN | 77.7% | 0.031 s |
| PCA-RBF NN | 79.3% | 0.023 s |
| PCA-ELM | 83.3% | 0.016 s |
| PCA-KELM | 84.7% | 0.016 s |
| PCA-ABC-ELM | 88.7% | 0.016 s |
| PCA-ABC-KELM | 90.3% | 0.016 s |

## 4. Conclusion

Network intrusion detection is very important for the computer security. This paper proposed a new intrusion detection method based on the ABC optimized KELM. The innovation of this work lies in the development and implementation of the PCA and ABC-KELM in the intrusion detection for the first time. Experimental tests have been carried out to evaluate the performance of the new method. The test result has showed satisfactory and effective intrusion detection performance of the proposed ABC-KELM method. In addition, through comparison between KELM, ELM, BP NN and RBF NN, it proves that the performance of the proposed PCA-ABC-KELM method is superior to its rivals in terms of both detection accuracy and training speed. Thus, the Proposed ABC-KELM method shows promising applications in the domain of network intrusion detection. Future research will focus on the industrial practice of the newly proposed method.

## References

[1] Zhu S, Hu B. Hybrid Feature Selection Based on Improved GA for the Intrusion Detection System. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(4): 1725–1730.
[2] He M, Qiu D. An Intrusion Detection Method Based on Neighborhood Rough Set. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(7): 3736–3741.
[3] Yu G, Weng K. Intrusion detection system and technology of layered wireless sensor network based on Agent. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(8): 4238–4243.
[4] Saeid A, Behzad Z, Ahmad H, Behzad B. Synthetic Feature Transformation with RBF neural network to improve the Intrusion Detection System Accuracy and Decrease Computational Costs. *International Journal of Information and Network Security*. 2012; 1(1): 28–36.
[5] Chandrashekhar A, Raghuveer K. Performance Evaluation of Data Clustering Techniques Using KDD Cup99 Intrusion Detection Data Set. *International Journal of Information and Network Security*. 2012; 1(4): 294–305.
[6] Paliwal M, KumarA U. Neural Networks and Statistical Techniques: A Review of Applications. *Expert Systems with Applications*. 2009; 36: 2–17.
[7] Zong W, Zhou H, Huang G, Lin Z. Face Recognition Based on Kernelized Extreme Learning Machine. *Lecture Notes in Computer Science.* 2011; 6752: 263–272.
[8] Huang G, Chen L. Enhanced Random Search Based Incremental Extreme Learning Machine. *Neurocomputing*. 2008; 71: 16–18.
[9] Zong W, Huang G, Lin Z. Face Recognition Based on Extreme Learning Machine. *Neurocomputing*. 2011; 74: 2541–2551.
[10] Deng W, Zheng Q, Zhang K. Reduced Kernel Extreme Learning Machine. *Advances in Intelligent Systems and Computing*. 2013; 226: 63–69.
[11] Karaboga D, Basturk B. A Powerful And Efficient Algorithm For Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm. *Journal of Global Optimization*. 2007; 39(3): 459–471.