TELKOMNIKA, Vol.12, No.4, December 2014, pp. 1088~1095
ISSN: 1693-6930, accredited A by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v12i4.889

A Novel Intrusion Detection Approach using Multi-Kernel Functions

Li Jiao Pan*, Weijian Jin, Jin Wu
School of Electro-Mechanical & Information Technology, Yiwu Industrial & Commercial College
No. 2 Xueyuan Road, Yiwu 322000, China, telp/fax 0579-83803612
*Corresponding author, e-mail: panlijiao_ywu@163.com

Abstract

Network intrusion detection finds variant applications in computer and network industry. How to achieve high intrusion detection accuracy and speed is still received considerable attentions in this field. To address this issue, this work presents a novel method that takes advantages of multi-kernel computation technique to realize speedy and precise network intrusion detection and isolation. In this new development the multi-kernel function based kernel direct discriminant analysis (MKDDA) and quantum particle swarm optimization (QPSO) optimized kernel extreme learning machine (KELM) were appropriately integrated and thus form a novel method with strong intrusion detection ability. The MKDDA herein was firstly employed to extract distinct features by projecting the original high dimensionality of the intrusion features into a low dimensionality space. A few distinct and efficient features were then selected out from the low dimensionality space. Secondly, the KELM was proposed to provide quick and accurate intrusion recognition on the extracted features. The only parameter need be determined in KELM is the neuron number of hidden layer. Literature review indicates that very limited work has addressed the optimization of this parameter. Hence, the QPSO was used for the first time to optimize the KELM parameter in this paper. Lastly, experiments have been implemented to verify the performance of the proposed method. The test results indicate that the proposed LLE-PSO-KELM method outperforms its rivals in terms of both recognition accuracy and speed. Thus, the proposed intrusion detection method has great practical importance.

Keywords*: network intrusion detection, multi-kernel function based kernel direct discriminant analysis, kernel extreme learning machine, quantum particle swarm optimization*

1. Introduction

Along with the rapid development of internet and the associated application networks, network security has become a prominent and tough problem, in particular, intrusions and attacks on computer network systems becomes more complex and diverse. Huge economic losses have been caused by the computer and network intrusions and attacks every year. Therefore, it is essential to detect the intrusions and attacks in time to prevent damages of computers and networks.

The diversity and the evolution of the intrusion viruses make it very difficult in detecting and identifying the undergoing network intrusion. From the early worm virus to the recent shock, shock waves and the panda incense viruses, the attacking objects almost include all computer system accessed to the internet. The attacking viruses will cost the system resources, manipulate data and steal the confidential information, leading to massive economic losses. Typical intrusion viruses including the Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R), and Probe or Scan (PoS). Beside, American business magazine "Information Weekly" has published a survey on the network intrusion and indicates that a network attack happens every second in the global scope. In such a situation, network security has become an urgent and practical problem and received worldwide attentions. How to develop and use the existing security technology to protect all kinds of resources from damage is the hot spot in the research field of network security. Effective intrusion detection technology is the key issue to solve this problem.

The machine learning is a very useful technology in the field of computer and network security. However, the network intrusions are always contaminated by background noise. In addition, the high dimensionality of the intrusion data increases their detection difficulties [1]. Hence, it is crucial to eliminate useless information and extract distinct features in a low

Received September 3, 2014; Revised November 5, 2014; Accepted November 20, 2014

dimensional space. By doing so, the computation cost and the detection performance can be simultaneously optimized. Although the principal component analysis (PCA) [2] and its derivative algorithms have been proven to be powerful for feature extraction in the low dimensional space, the limitation is that the PCA cannot extract nonlinear properties of the original data [3]. In contrast to PCA, the KDDA can extract nonlinear properties from the original data [4]. The KDDA adopt the kernel trick to reduce the high dimensional image data into a much lower dimension space by keeping the nonlinear properties of the original data. By doing so, the nonlinear properties of the data of interest can be obtained. The construction of the kernel function greatly determines the performance of KDDA in feature extraction. In most existing kernel functions of KDDA, single kernel was used; however, recent research result shown great interest in multi-kernel. Multiply kernels would have more excellent characteristics than single kernel and thus provide better performance of KDDA. Literature review indicates that the limited work has been done to address the multi-kernel issue for KDDA in the intrusion detection [5]. Hence, the outcomes of the multi-kernel function based KDDA (MKDDA) should be evaluated.

On the other hand, the artificial intelligence has been extensively used in the network intrusion detection, such as artificial neural network (ANN) and support vector machine (SVM) [6-9]. However, the ANNs, including BP NN and RBF NN, are often suffer from local minima and slow convergence speed [6]; and the SVM needs to determine the kernel function, error control parameters, and penalty coefficient. Hence, although ANN and SVM have a lot of advantages in machine learning [10], they face challenges on learning speed and scalability, which limit their applications in network intrusion detection. In order to overcome this problem, the kernel extreme learning machine (KELM) has been proposed as en integration of ANN and SVM to provide quick and accurate pattern recognition ability [11]. The KELM has the advantages of both ANN and SVM while only needs to set up one only parameter, i.e. the number of hidden layer nodes of the network [11]. Zong and Huang [11] have presented the KELM in the face recognition and found that the KELM outperforms LS-SVM in terms of both recognition prediction accuracy and training speed. However, a parameter optimization mechanism of the KELM has not well developed in existing work. Proper setting of the neuron number of the KELM can enhance the training speed and accuracy [12, 13]. It is therefore imperative to develop an optimization mechanism for the KELM.

To enhance the network intrusion detection, this work presents a new method based on the MKDDA and QPSO-KELM. Compared with PCA, the proposed method has employed the MKDDA to extract nonlinear features of the face images. It also developed the optimized KELM for faster and more precise intrusion detection when comparing with ANN and SVM. Experimental analysis has verified high performance of the proposed method.

2. Research Method

In this work, the intrusion detection method based on the MKDDA and QPSO-KELM has been proposed. A brief description about the proposed method is illustrated as follows.

2.1. MKDDA

Assume $\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_m] \in R^p$ and its subset $\mathbf{X}_i \in \mathbf{X}$ with *q* elements. Let $\Phi : x \in R^N \to \Phi(x) \in \mathbf{F}$ a nonlinear mapping from input space to a high dimensional feature space **F**, in which the inner-class and inter-class scatter matrices are $\mathbf{S}_w$ and $\mathbf{S}_b$. Then we define $\mathbf{S}_w$ and $\mathbf{S}_b$ as follows:

$$S_b = \frac{1}{p}\sum_{i=1}^{m} q_i (\bar{\Phi}_i - \bar{\Phi})(\bar{\Phi}_i - \bar{\Phi})^{\mathrm{T}}, \tag{1}$$

$$S_w = \frac{1}{p}\sum_{i=1}^{m}\sum_{j=1}^{q} (\bar{\Phi}_{ij} - \bar{\Phi}_i)(\bar{\Phi}_{ij} - \bar{\Phi}_i)^{\mathrm{T}}, \tag{2}$$

Where, $\Phi_{ij} = \Phi(x_{ij})$; $\Phi_i = \frac{1}{q}\sum_{j=1}^{q}\Phi(x_{ij})$ denotes the sample mean of class $\mathbf{X}_i$; $\Phi = \frac{1}{p}\sum_{i=1}^{p}\sum_{j=1}^{q}\Phi(x_{ij})$ denotes the average of the sample. Then we could obtain

$$S_b = \sum_{i=1}^{m}(\sqrt{\frac{q_i}{p}}(\bar{\Phi}_i - \bar{\Phi}))(\sqrt{\frac{q_i}{p}}(\bar{\Phi}_i - \bar{\Phi}))^{\mathrm{T}} = \sum_{i=1}^{m}\bar{\bar{\Phi}}_i\bar{\bar{\Phi}}_i^{\mathrm{T}} = \Phi_b\Phi_b^{\mathrm{T}}, \quad (3)$$

Where, $\bar{\bar{\Phi}}_i = \sqrt{\frac{q_i}{p}}(\bar{\Phi}_i - \bar{\Phi})$ and $\Phi_b = [\bar{\bar{\Phi}}_1 \ldots \quad \bar{\bar{\Phi}}_m]$. Form (3) we can see that the dot product is required to calculate $\Phi_b\Phi_b^{\mathrm{T}}$. The dot product is very computational cost and in order to avoid it the kernel function has been proposed to compute $\Phi_b\Phi_b^{\mathrm{T}}$ [4].

In this work, we proposed a multi-kernel function that integrated the radial basis function (RBF) kernel and polynomial kernel functions to provide more efficient kernel function computation [4]. The multi-kernel function is described as

$$K(x,y) = \alpha\exp(-\frac{\|x-y\|^2}{2\sigma^2}) + (1-\alpha)(x\cdot y + b), \quad (4)$$

Where, x, y are the inputs, α, σ and b are constants. Then, the $\Phi_b\Phi_b^{\mathrm{T}}$ can be calculated by the kernel function rather than the dot production. Thus, the original dataset **X** could be projected into low dimentional feature space **F** by solve the eigenvalue problem [4].

2.2. QPSO-KELM

Given samples $\{(y_i, g_i) : i = 1, 2, ..., N; y_i \in R^p, g_i \in R^q\}$, where *y* is the feature vector and *g* is the class label vector, the below function is used to identify the sample [10]

$$\sum_{i=1}^{n}\xi_i \mathrm{s}(\varsigma_i^T y_j - b_i) = o_i, j=1, 2,..., N. \quad (5)$$

Where, $\mathrm{s}(\cdot)$ is the activation function; *n* is the number of hidden neuron; o_i is the output of *j*th sample; ς_i and ξ_i are the input and output weight vectors; b_i is the threshold of the *i*th hidden neuron. If the output o can approximate *g*, then

$$\sum_{i=1}^{n}\xi_i \mathrm{s}(\varsigma_i^T y_j - b_i) = o_j = t_j, j=1, 2 ,..., \mathrm{N}. \quad (6)$$

Hence, we derive

$$\mathbf{H}\boldsymbol{\xi} = \mathbf{T}, \quad (7)$$

Where,

$$\mathrm{H} = \begin{bmatrix} \mathrm{s}(\varsigma_1^T y_1 - b_1) & \cdots & \mathrm{s}(\varsigma_n^T y_1 - b_n) \\ \vdots & \cdots & \vdots \\ \mathrm{s}(\varsigma_1^T y_N - b_1) & \cdots & \mathrm{s}(\varsigma_n^T y_N - b_n) \end{bmatrix},$$

$$\boldsymbol{\xi} = [\xi_1, \cdots, \xi_n]^T \text{ and } \mathbf{G} = [g_1, \cdots, g_N]^T.$$

To solve (7), the KELM adopts a least squares error to get solution:

$$\overline{\boldsymbol{\xi}} = \mathbf{H}^{\dagger}\mathbf{T}, \tag{8}$$

where, $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of $\mathbf{H}$. Function $\mathrm{s}(\cdot)$ is usually unknown, we replace it by the kernel matrix $\mathbf{K} = \left[\mathrm{K}(x; x_1) \quad \cdots \quad \mathrm{K}(x; x_N)\right]^T$ ($\mathrm{K}(\cdot)$ is the kernel function). Then it yields

$$\mathbf{o} = \mathbf{KT} \tag{9}$$

Herein, the Gaussian kernel function (RBF) is adopted. The number of hidden neuron *n* is difficult to determine. Hence, QPSO was used to obtain a proper *n* [11].

2.3. The Proposed Intrusion Detection Method

In this paper the novel development using MKDDA-QPSO-KELM are proposed for the network intrusion detection. The proposed network intrusion detection processes are given as follows:

Step 1: Pre-treat the original network intrusion data to standardized data format.

Step 2: Extract distinct features from the input network intrusion data in the form of manifold by MKDDA.

Step 3: Train the KELM using the new features, and determine the neuron number of hidden layer of KELM using QPSO.

Step 4: Test the performance of the proposed network intrusion detection model, and provide the test result as the base for a valid network intrusion management decision. A diagram block of the proposed network intrusion detection method is illustrated in Figure 1.

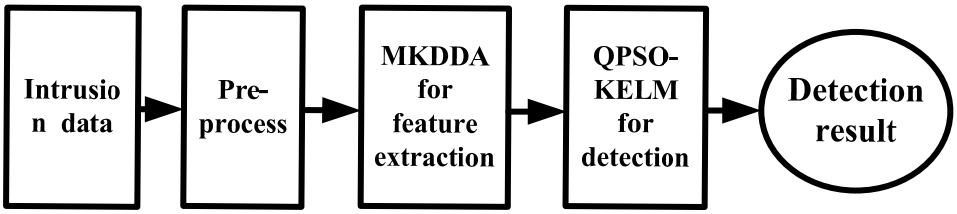

Figure 1. The proposed network intrusion detection method

3. Results and Analysis

In order to evaluate the performance of the proposed computer intrusion method, experiment tests have been implemented in this work. Fig. 2 shows the experiment set-up. A mini network was established by using one linux server and one windows server, as well as three windows hosts and three linux hosts. The Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R) and Probe or Scan (PoS) were simulated using this experiment set-up. We have collected 3,000 samples for each intrusion type and 35 features for each sample. These features include the bytes issued from source to destination, the bytes from destination to source, duration, teardrop, neptune, etc. Herein 5, 000 samples of each intrusion type have been recorded for the experimental test. Figs. 3~5 show the feature selection results by three popular methods, i.e. Kernel PCA (KPCA), KDDA and MKDDA.

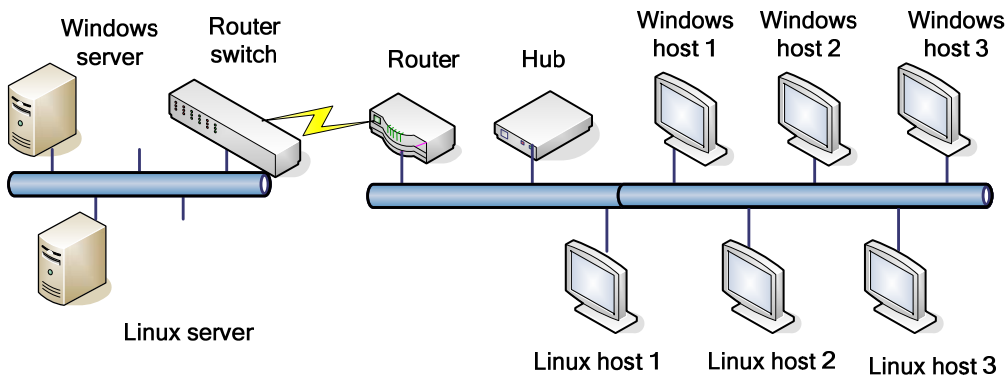

Figure 2. The experiment set-up

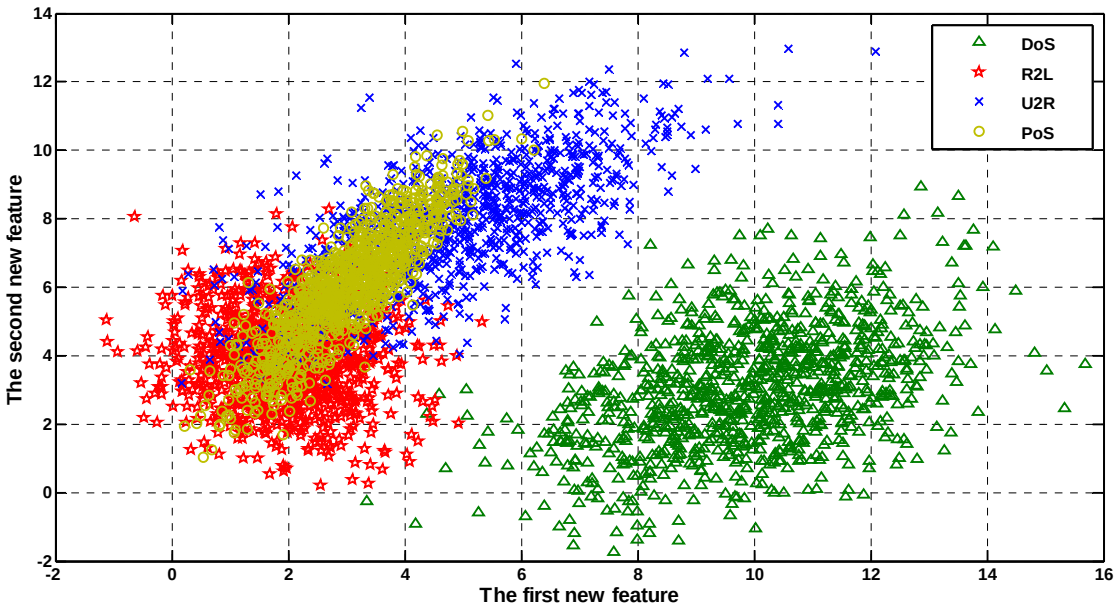

Figure 3. Feature extraction result using KPCA

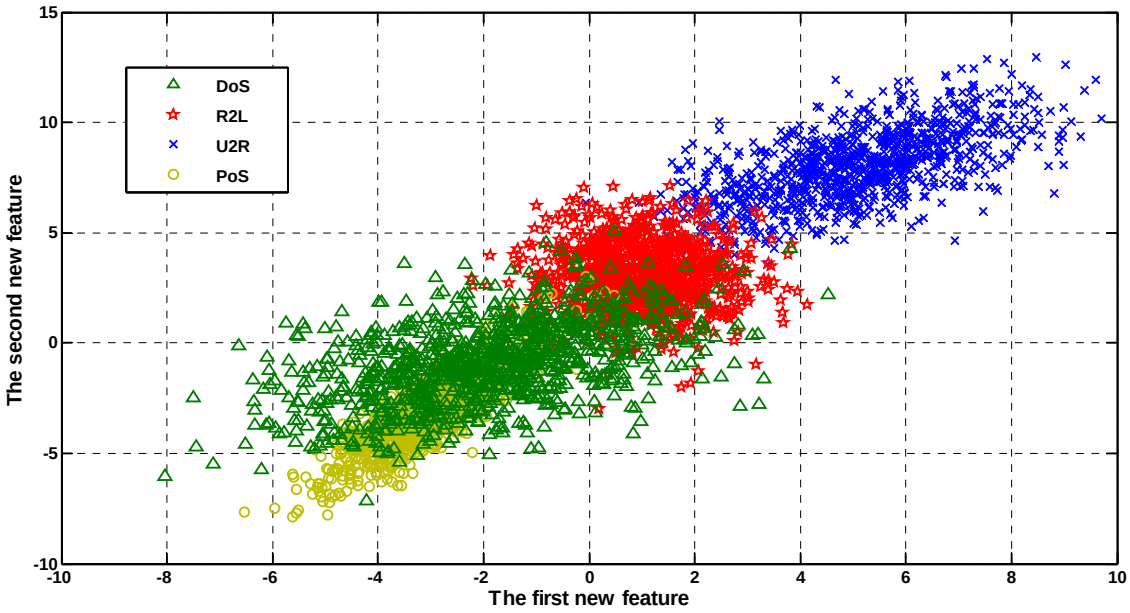

Figure 4. Feature extraction result using KDDA

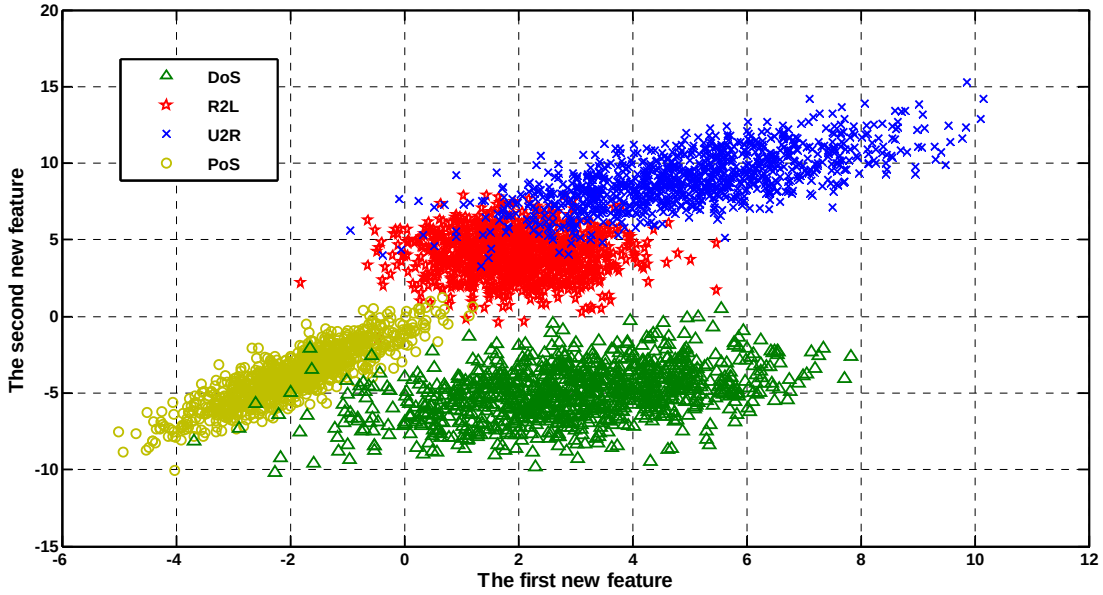

Figure 5. Feature extraction result using MKDDA

It can be seen in Figure 3 that when using KPAC to select the most useful features just the intrusion type of DoS could be identified well while the other three types were overlapped with each other. This means the feature selection rate of the KPCA is very low in this study. In Fig. 4 we could note that the KDDA can separate two types of intrusions, i.e. R2L and U2R; however, the DoS and PoS were completely mixed up. In contrast, when the MKDDA was adopted in Fig. 5 it indicated that the four types of intrusions had been well recognized by clear boundaries though a small portion of the intrusions data mixed together. As a result, the feature selection performance of the MKDDA was superior than that of KPCA and KDDA.

Table 1 lists the comparison of the proposed method against some existing approaches.

Table 1. The comparison results of the face recognition

Method	Detection rate (%)	Computation time (s)	Method	Detection rate (%)	Computation time (s)
PCA-SVM	77.7	5.3	KPCA-SVM	78.7	5.3
PCA-ANN	75.3	5.4	KPCA-ANN	75.7	5.5
PCA-KELM	77.7	4.9	KPCA-KELM	79.3	5.0
PCA-QPSO-KELM	78.7	4.5	KPCA-QPSO-KELM	79.7	4.5
KDDA-SVM	83.3	5.2	MKDDA-SVM	83.7	5.2
KDDA-ANN	83.3	5.2	MKDDA-ANN	83.7	5.3
KDDA-KELM	83.7	4.9	MKDDA-KELM	84.3	4.9
KDDA-QPSO-KELM	84.3	4.5	MKDDA-QPSO-KELM	85.3	4.4

In Table 1 it can be seen that the MKDDA-QPSO-KELM outperformed the other methods and obtained the best intrusion detection rate. The best detection rate is 85.3%, which is 1.0% higher or more than the SVM and ANN based approched. This is because KELM has advantages of both SVM and ANN and in the same time KELM only uses one parameter to lighten its structure to obtain better learning ability than SVM and ANN [7]. One also can note that the intrusion detection rate of MKDDA-QPSO-KELM is 1.0% higher than MKDDA-KELM. This means that the QPSO has contributed some efford to enhance the KELM recognition performance. By applying QPSO to KELM, its neuron number could be optimized to yield fater and higher recognition rate. This explains why the proposed method generated better intrusion detection rate than the others. Hence, the comparison results indicate that the KELM has fast training speed owning to its special structure, and the MKDDA-QPSO-KELM can improve the intrusion detection rate.

4. Conclusion

In order to develop an efficient method for intrusion detection, a new method based on MKDDA and QPSO-KELM has been presented to emhance the detection accuracy and computation speed in this work. The inovation of the proposed work is that the multi-kernel function based KDDA was proposed to give better feature selection performace than a single kernel function of KDDA. Meanwhile, the QPSO was employed to midify the only parameter of the KELM and hence reasonable KELM structure could be obtained to enhance the intrusion detection rate. Experimental tests have been carried out to verify the proposed method. The analysis results demonstrate that good intrusion detection performance could be attained by the proposed method. In addition, through comparison between different feature extraction algorithms (i.e. PCA, KPCA, KDDA and MKDDA) and different intelligent separators (i.e. SVM, ANN and KELM) the proposed MKDDA-QPSO-KELM method generated the best performance in terms of accuracy and computation cost. Thus, the proposed method has practical importance. Future research will focus on the industrial practice of the proposed method.

References

[1] Zhu S, Hu B. Hybrid feature selection based on improved GA for the intrusion detection system. *TELKOMNIKA*. 2012; 11(4): 1725–1730.
[2] Hou G, Ma X, Zhang Y. A new method for intrusion detection using manifold learning algorithm. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(12): 7344–7350.
[3] 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.
[4] Li F, Tang W, Duan H, Hao J. Application of fractional power polynomial kernel function to kernel direct discriminant analysis. *Optics and Precision Engineering*. 2007; 15(9): 1140-1144.
[5] Li F, Xu K. Kernel model applied in kernel direct discriminant analysis for the recognition of face with nonlinear variations. *Transactions of Tianjin University*. 2006; 12(2): 147-152.
[6] Liang L. Network intrusion detection system based on optimized Fuzzy rules algorithm. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(4): 2816–2825.
[7] Chen S, Diao H. A network intrusion detection method based on improved ACBM. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(4): 2808–2815.
[8] Liu L, Wan P, Wang Y, Liu S. Clustering and hybrid genetic algorithm based intrusion detection strategy. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(1): 762–770.

[9] 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): 4238–4343.
[10] Huang G, Chen L. Enhanced random search based incremental extreme learning machine. *Neurocomputing.* 2008; 71: 16–18.
[11] Zong W, Huang G, Lin Z. Face recognition based on extreme learning machine. *Neurocomputing*. 2011; 74: 2541–2551.
[12] Jiang Y, Wu J, Zong C. An effective diagnosis method for single and multiple defects detection in gearbox based on nonlinear feature selection and kernel-based extreme learning machine. *Journal of Vibroengineering*. 2014; 16(1): 499–512.
[13] Sun H, Qi Y. A chaos cloud particle swarm algorithm based available transfer capability. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(1): 38–47.