Feature Extraction based Approaches for Improving the Performance of Intrusion Detection Systems

Long-Sheng Chen*, Jhih-Siang Syu

Abstract—**In recent years, the rapid development of information and communication technology results in too many loopholes in the network, and thus attracts lots of hackers' attacks. Intrusion Detection System (IDS) has been developed to detect these attacks. Depending on different data and analysis methods, this system will have different detection methods. But, there is no one model is absolutely effective. Therefore, this study will focus on improving the classification performance of anomaly detection. In this study, we'll propose "Local Latent Semantic Indexing (LLSI)" and "Local Kernel-Principal Component Analysis (LKPCA)" based methods, which introduce class information to feature extraction techniques. And the proposed methods will be integrated into support vector machine (SVM) to improve the performance of classification. Finally, KDD-NSL data set will be employed to testify the effectiveness of the proposed methods.**

Index Terms—**Intrusion Detection System, Feature Extraction, Latent Semantic Indexing, Kernel-Principal Component Analysis, Dimension Reduction.**

I. Introduction

Information and communication technology (ICT) has become an indispensable part of human life based on well-built infrastructure. No matter government, business or a variety of academic, medical, and other organizations, they increasingly rely on ICT. But, it also brings lots of security problems and crisis. Lots of network attack tools can easily be found and downloaded on the Internet. Through a variety of network vulnerabilities and continuously developed new attack techniques and tools, it leads to cyber-attacks continue to evolve. So, this problem cannot be underestimated (Feng et al, 2014; Hubballi and Suryanarayanan, 2014; Govindarajan and Chandrasekaran, 2012).

To solve this problem, intrusion detection systems (IDS) have been firstly developed by Anderson (1980). Since then, various attack identification techniques have been proposed, such as rule-based, neural networks, support vector machines, and so on. Lots of well-constructed intrusion detection systems have been developed and have been applied to real world computer systems (Mohammed and Sulaiman, 2012). But, 30 years has passed since Anderson's first study. We cannot complete solve this problem, mainly because of dramatically advances in technology.

Depending on data type and data analysis methods, IDS could be divided into several different modes. Based on analytical methods, IDS can be divided into misuse detection and anomaly detection. The former uses a known attack signature database, whenever the login data match any feature, IDS will give alarms. The latter uses normative behavior model and then observes behaviors' deviation. If the behavior is inconsistent with the model, IDS will give alarms. The accuracy of misuse detection system is high, but it cannot detect unknown attacks. Anomaly detection methods have high false alarm rate, but it's able to detect unknown attacks (Hubballi and Suryanarayanan, 2014). So, no single one intrusion detection system is applicable to any situation.

In addition to a variety of algorithms and intrusion detection system model, dimension reduction methods are often used to select important features and to reduce dimension size for saving computational cost. Typically, there are two groups of algorithms to represent the feature space used in classification. The first one is feature selection which is to select a subset of most representative features from the original feature space. The second algorithm is feature extraction which is to transform the original feature space to a smaller feature space to reduce the dimension. These dimension reduction techniques have widely applied to solve real problems. For examples, Eesa et al. (2015) used cuttlefish based feature selection techniques to maintain data quality features and remove redundant and irrelevant features. Gan et al. (2013) combined with Partial Least Square (PLS) feature extraction technique and Core Vector Machine (CVM) algorithm to detect the abnormalities. Kuang et al. (2014) using a support vector machine (SVM) combined kernel principal component analysis (KPCA) and genetic algorithm (GA), where KPCA is used to reduce the data dimension.

To sum up, feature extraction which is one of dimension reduction method can produce a new set of features by transforming the original input variables (Yang et al., 2011). But, the new set of features cannot retain the original meanings of original features (Jain et al., 2000). Liu et al. (2004) and Chen et al. (2008) indicated that feature extraction can reduce the dimensions of the feature space greatly compared with feature selection.

Manuscript received January 8, 2015; accepted January 22, 2015. This work was supported in part by the National Science Council of Taiwan, R.O.C. (Grant No. MOST 101-2628-E-324-004-MY3).

*L.-S. Chen is an Associate Professor of Department of Information Management and Dean of Student Affairs Office with the Chaoyang University of Technology, Taichung 41349, Taiwan (phone: 886-4-23323000ext7752; fax: 886-4-23304902; e-mail: lschen@cyut.edu.tw).

J.-S. Syu is a graduate student with the Department of Information Management, Chaoyang University of Technology, Taichung 41349, Taiwan (e-mail: s10314609@gm.cyut.edu.tw).

Because the range of intrusion detection system is wide, the major objective of this study is to focus on improving the classification performance of anomaly detection. Latent Semantic Indexing (LSI) (Deerwester et al., 1990; Guan et al., 2013; Uysal and Gunal, 2012; Wang et al., 2015) and KPCA will be employed to integrate into Support Vector Machines (SVM) to increase the anomaly detection performance. In this study, we'll to propose a "Local Latent Semantic Indexing (LLSI) by singular value decomposition (SVD)" and Local Kernel-Principal Component Analysis (LKPCA)" based methods which introduce class information to feature extraction techniques. And the proposed methods will be integrated into support vector machine (SVM) to improve the performance of classification. Finally, KDD-NSL data set will be employed to testify the effectiveness of the proposed methods.

II. Literature Review

A. Intrusion Detection Systems

Those unauthorized activities which have been designed to access system resources or data are called "intrusion" (Hubballi and Suryanarayanan, 2014). The main core of intrusion detection system is to be able to detect these attacks or illegal activities to provide network managers corresponding treatment.

The IDS can be classified into several different modes according to its data source and analytic methods. Table 1 shows the category of IDS. Usually, we build IDS in the front end of a network segment, or behind/after the firewall, to analyze packet through the inspected network section.

In 30-year history of IDS, rule-based early detection module has been firstly developed and become the main stream (Han, 2003; Lee, 1999). After then, different algorithms based detection methods, such as genetic algorithm (GA) (Kuang et al., 2014), Bayes (Benferhat et al., 2013), neural networks (NN) (Thomas and Balakrishnan, 2008), and support vector machine (SVM) (Feng et al, 2014; Mohammed and Sulaiman, 2012) have been constructed.

TABLE I
IDS CLASSIFICATION

Classification	Species	Monitoring objectives and approach.
Sources of information to distinguish	Network-based	Monitoring network packets.
	Host-based	Internal records monitoring activities of the host system.
	Application-based	Log file monitoring application generated.
	Goal-based	Monitoring special or secret archives.
	Hybrid	Monitoring network packets and host systems combined record of activities.
Analytical methods to distinguish	Misuse-based	Uses a database of known attack signatures, whenever the login data match any feature, IDS will give alarms.
	Anomaly-based	Normative system behavior and observable deviations are raised as alarms.
	Hybrid	Combines signature-based and anomaly-based approach enables them to complement each other.

Till now, in order to cope with the rapid development of information technology, single one type of models has been insufficient to protect network security. So, the hybrid approaches gradually become the mainstream (Kim and Kim, 2014). "False positives" (FP) is the most common indicators in assessing the quality of an IDS. And, how to reduce the so-called false positive rate in this field has become one of important issues (Hubballi and Suryanarayanan, 2014).

However, the real purpose of this study is not to build complete IDS, but how to improve the classification performance. Kuang et al. (2014) pointed out that the intrusion detection can be seen as essentially a classification problem to distinguish abnormal activities. Thus, the study focuses on the emphasis on the use of data mining techniques to detect abnormal patterns of classification performance, hoping to improve their performance through effective data mining techniques.

B. Feature Extraction

According to available literatures, lots of works used various feature extraction methods for dimension reduction (Hoque and Bhattacharyya, 2014; Lin et al., 2012). The most representative feature extraction algorithm is the Latent Semantic Indexing (LSI) (Deerwester et al., 1990; Guan et al., 2013; Uysal and Gunal, 2012; Wang et al., 2015) which is an automatic method that transforms the original textual data to a smaller semantic space by taking advantage of some of the implicit higher-order structure in associations of words with text objects (Berry et al., 1995; Deerwester et al., 1990). The transformation is computed by applying truncated singular value decomposition (SVD) to the term-by-document matrix. After SVD, terms which are used in similar contexts will be merged together.

SVD is an optimal linear transformation for dimensionality reduction. It allows the arrangement of the space to reflect the major associative patterns in the data, and ignore the smaller, less important influences. SVD transformation as well has the advantage of yielding zero-mean and uncorrelated features (Castelli et al., 2003). Moreover, it has been reported that SVD can be applied to education, solving linear least-squares problems, data compression (Akritas and Malaschonok, 2004), document classification (Guan et al., 2013), and text classification (Uysal and Gunal, 2012). Therefore, LSI is employed as the feature extraction tool in this study.

Another famous feature extraction method is Principal Component Analysis (PCA). PCA method can only extract the linear structure information in the data set, however, it cannot extract this nonlinear structure information. Except intrusion detection, there are lots of successful applications in many areas, such as face recognition, stock prediction model, and so on (Zhou et al, 2014; Kuang et al, 2014; Wang and Battiti, 2006;Wen et al. 2014).

KPCA (Scholkopf et al., 1998) is an improved PCA, which extracts the principal components by adopting a nonlinear kernel method. A key insight behind KPCA is to transform the input data into a high dimensional feature space *F* in which PCA is carried out, and in implementation, the implicit feature vector in *F* does not need to be computed explicitly, while it is just done by computing the inner product of two vectors in *F* with a kernel function (Kuang et al., 2014).

Using non-linear kernel function, KPCA improve the nonlinear problems which cannot be solved by PCA (Chen et al, 2008; Ding et al, 2009). The main advantage of KPCA is that it does not involve nonlinear optimization; essentially it only requires linear algebra, which makes it as simple as standard PCA (Jia et al., 2012). KPCA requires only the solution of an eigenvalue problem, and due to its ability to use different kernels, it can handle a wide range of nonlinearities. In addition, KPCA does not require the number of components to be extracted and specified prior to modeling. Due to these advantages, this study employs KPCA and compare its results to LSI.

III. Implemental Procedure

The experimental procedure is divided into two stages, the first stage, we use feature extraction methods to reduce dimensionality. The original 41 features will be reduced to smaller dimensionality. In the second phase, we find the optimal reduced dimension size through evaluating by SVM classifier. The implemental procedure can be shown in Figure 1. Actually, there are 6 major steps. The concise steps can be found as follows.

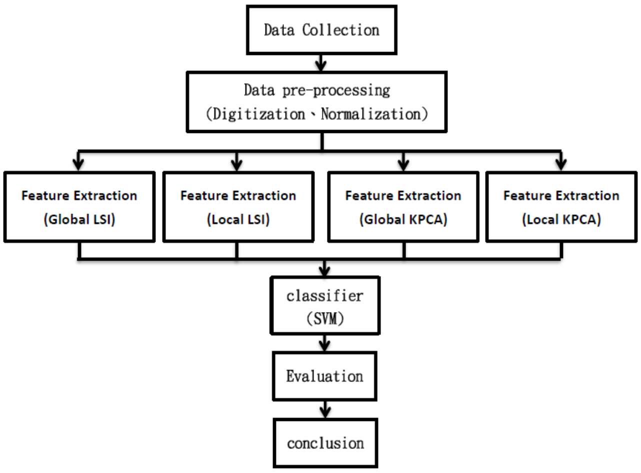
Data Collection
Data pre-processing
(Digitization、Normalization)
Feature Extraction (Global LSI)
Feature Extraction (Local LSI)
Feature Extraction (Global KPCA)
Feature Extraction (Local KPCA)
classifier (SVM)
Evaluation
conclusion

Figure 1 The implemental procedure of this work

Step 1: Data Collection

The employed “train20percent” data comes from KDD-NSL data sets (Nsl-kdd, 2009). This data set is used to detect intrusion, and provided by well-known public data sets KDD CUP’ 99 (Hettich and Bay, 1999). It’s modified from KDD CUP’ 99, and improve some disadvantages of original data (Tavallaee et al., 2009).

Step 2: Data Preprocessing

In train20percent dataset, the attack type has been categorized 23 types. We combine all 23 attacks into one class “abnormal”. Therefore, it has become binary class classification problems.

Step 3: Feature Extraction

This study uses 4 feature extraction methods, Global LSI, Local LSI, Global KPCA, and Local KPCA. Global LSI and Global KPCA mean the original LSI and KPCA without introducing class information. Local LSI and Local KPCA are our presented methods.

Step 3.1 LSI

Let’s briefly introduce the concept of SVD. Let *A* be a $m \times n$ matrix of rank *r* whose rows represent documents and columns denote terms (variables). Let the singular values of A (the Eigen values of $A \times A^T$) be $\sigma_1 \ge \sigma_2 \ge ...... \ge \sigma_r$. The *singular value decomposition* of *A* expresses *A* as the product of three matrices $A = USV^T$, where $S = diag(\sigma_1,..., \sigma_r)$ is an $r \times r$ matrix, $U = (u_1,..., u_r)$ is an $m \times r$ matrix whose columns are orthonormal, and $V^T = (v_1,..., v_r)^T$ is an $r \times n$ matrix. LSI works by omitting all but the *k* largest singular values in the above decomposition, for some suitable *k* (*k* is the dimension of the low-dimensional space). It should be small enough to enable fast retrieval and large enough to adequately capture the structure of the corpus. Let $S_k = diag(\sigma_1,..., \sigma_k)$, $U_k = (u_1,..., u_k)$ and $V_k = (v_1,..., v_k)$. Then $A_k = U_k S_k V_k^T$ is a matrix of rank *k*, which is the approximation of A. The rows of $V_k S_k$ above are then used to represent the documents. In other words, the row vectors of *A* are projected to the *k*-dimensional space spanned by the row vectors of U_k; we sometimes call this space the LSI space of *A*.

We implement Global LSI and Local LSI. Global LSI is the general method of using SVD. We can choose the reduced dimension size k, and then we use $M_k = U_k \times S_k$ to be a new set of input features. In Local LSI (Liu et al., 2004), we first divide data into several groups based on their class labels. Next, we implement the same procedure with Global LSI. So, Local LSI introduces the additional information of class (dependent variables) while transforming.

Step 3.2 KPCA

Global KPCA

Step 3.2.1 Collect data $X_{n \times m}$, and normalize the data of each variable into mean 0 and variance 1.

Step 3.2.2 Compute the kernel matrix $K \in R^{I \times I}$, note the elements as K_{ij}.

Step 3.2.3 Carry out centering in the feature space for K.

Step 3.2.4 Carry out principal component decomposition for *k*, and determine the number of PCs retained, recorded as A, and then projection is obtained.

Local KPCA

Step 3.3.1 Collect data $X_{n \times m}$, and normalize the data of each variable into mean 0 and variance 1.

Step 3.3.2 Divide collected data into several groups based on their class labels. For each group, we implement sub-steps 3.3.3~3.3.5, respectively.

Step 3.3.3 Compute the kernel matrix $K \in R^{I \times I}$, note the elements as K_{ij}.

Step 3.3.4 Carry out centering in the feature space for K.

Step 3.3.5 Carry out principal component decomposition for *k*, and determine the number of PCs retained, recorded as A, and then projection is obtained.

Step 4 Build SVM classifier

In order to confirm the performances of proposed Local LSI and Local KPCA, we use the reduced dimensionality to build SVM. In this step, we use the training data to construct SVM classifier, and then input the test data to validate the built classifiers. Moreover, 5 fold cross validation experiment has been employed for these training data.

Step 5 Results Evaluation

We use overall accuracy (OA) and F1 to evaluate the performances.

Step 6 Draw Conclusions

Based on results, we can make conclusion.

IV. Results

A. Data Preprocessing

In this study, we employ train20percent file from NSL-KDD (Nsl-kdd, 2009). In this dataset, the attack type has been categorized in Table 2. Besides, Table 3 shows the data size and class distribution information. By the way, we define 4 attack types into one “abnormal” class. Therefore, it has become binary class classification problems. In addition, 5 fold cross validation experiment has been employed. All data will be normalized.

TABLE II
ATTACK TYPES IN NSL-KDD DATASET

Attack Type	Attack detailed classification
U2R	Buffer_overflow, loadmodule, multihop, perl, rootkit
R2L	ftp_write, guess_passwd, imap, phf, spy, warezclient, warezmaster
DOS	ftp_write, guess_passwd, imap, phf, spy, warezclient, warezmaster
Probe	Ipsweep, nmap, portsweep, satan

TABLE III
DATA DISTRIBUTION

Experiment No.	Data size	Class distribution
Fold-1	25,192	Normal：13,449 Attack：11,743
Fold-2		
Fold-3		
Fold-4		
Fold-5		

B. Measurement Index

To illustrate measurement index, we use table 4 to demonstrate accuracy and F1.

TABLE IV
BINARY CLASSIFICATION

	Predicted Normal	Predicted Attack
Actual Normal	TP	FP
Actual Attack	FN	TN

In Table 4, the meanings of denotations TP, FP, FN, TN have given as follows.

(1) TP: Actual normal examples classified into “normal”.
(2) FP: Actual normal examples classified into “attack”.
(3) FN: Actual attack examples classified into “normal”.
(4) TN: Actual attacks examples classified into “attack”.

$$OA = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

$$\Pr ecision = \frac{TP}{TP + FP} \quad (2)$$

$$\operatorname{Re} call = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 * \Pr ecision * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \quad (4)$$

The overall accuracy (OA) has been defined in equation (1). F1 is a weighted index both considering Precision and Recall indicators. Precision, Recall, F1 can be defined in equations (2)~(4).

C. Experimental Results

Figure 2 provides the summary of results when using LSI. From this figure, we can find that the performance of Global LSI and Local LSI keep stable when dimension size decreasing. But, even when the dimension size reduce from original 41 attributes to 1 attributes, Local LSI has better performance than Global LSI no matter considering OA or F1.

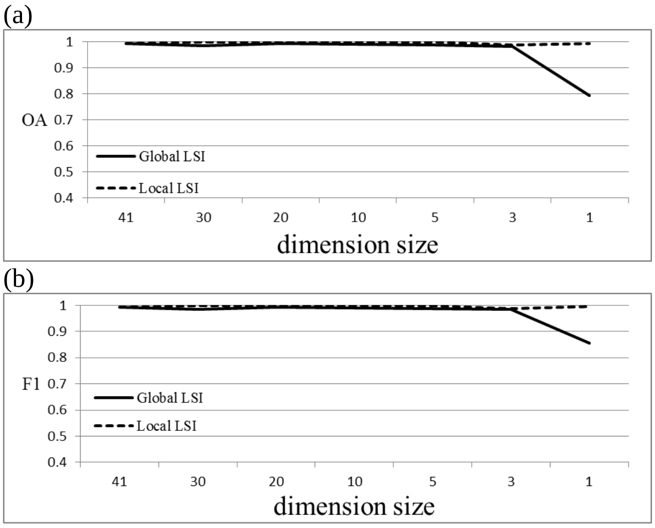

FIGURE 2 RESULTS OF GLOBAL LSI AND LOCAL LSI

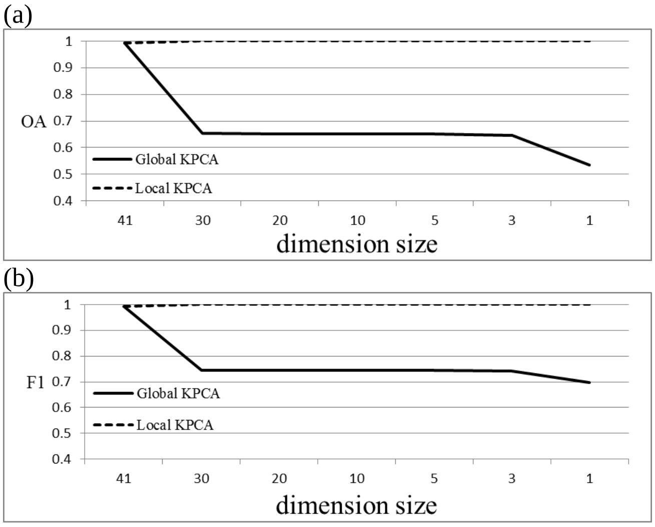

FIGURE 3 RESULTS OF GLOBAL KPCA AND LOCAL KPCA

Figure 3 provides the summary of results when using KPCA. From this figure, unlike LSI, we can find that the performance of Global KPCA gets worse when dimension size decreasing. But, even when the dimension size reduce from original 41 attributes to 1 attributes, Local KPCA significantly outperform Global LSI no matter considering OA or F1.

Table 5 gives the comparison of original SVM, Global

LSI, and Local LSI. From this table, we can find the computational time save up to 98% no matter implementing Global LSI or Local LSI. But, Local LSI can maintain the classification performance.

Table 6 shows the comparisons of original SVM, Global KPCA, and Local KPCA. Considering computational time, Global KPCA and Local KPCA only use 1.34% and 0.017%, respectively. Local KPCA significantly outperform original SVM no matter in OA and F1. From Tables 5~6, we can find that introducing class information to feature extraction methods can not only keep the classification performance, but also remarkably reduce the computational time.

TABLE V
COMPARISONS OF ORIGINAL SVM, GLOBAL LSI AND LOCAL LSI

	SVM-Original (dimensions 41)	Global LSI-SVM (dimensions 1)	Local LSI-SVM (dimensions 1)
OA (%)	99.31 (0.06)	79.28 (23.90)	99.39 (0.17)
F1 (%)	99.35 (0.05)	85.65 (15.38)	99.43 (0.16)
Time	345,643.78 (12.25)	7,244.82 (29.62)	7,202.11 (0.15)

Note: The number 99.31 (0.06) in this table means Mean (Standard Deviation), respectively.

TABLE VI
COMPARISONS OF ORIGINAL SVM, GLOBAL KPCA AND LOCAL KPCA

	SVM-Original (dimensions 41)	SVM-Global KPCA (dimensions 1)	SVM-Local KPCA (dimensions 1)
	Averages (Standard Deviation)		
OA (%)	99.31 (0.06)	53.39 (0)	100 (0)
F1 (%)	99.35 (0.05)	69.61 (0)	100 (0)
Time	345,643.78 (12.25)	4,646.39 (19.19)	60.4 (10.00)

Note: The number 99.31 (0.06) in this table means Mean (Standard Deviation), respectively.

V. CONCLUSIONS

In the present study, we utilize feature extraction selection method (LSI, KPCA) to improve the performance of intrusion detection systems. By introducing class information, we present Local-LSI and Local-KPCA. Results indicated that Local-LSI and Local-KPCA outperform Global LSI and Global KPCA, and original SVM, respectively. Local-KPCA cannot only dramatically save computational time, but also remarkably increase the classification performance. We can conclude that introducing class information to feature extraction methods can not only keep the classification performance, but also remarkably reduce the computational time.

However, the results obtained from one data set. If we want to have a generalized conclusion, more data sets and additional feature extraction methods can be used in the future works. Multi-class classification might be another one direction of future researches.

REFERENCES

[1] A. S. Eesa, Z. Orman, A. M. A. Brifcani, A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems, Expert Systems with Applications 42 (2015) 2670﹣2679.

[2] B. Scholkopf, A. Smola, K. R. Muller, Nonlinear Component Analysis as a Kernel Eigenvalue Problem, Neural Computation 10(5) (1998) 1299﹣1319.

[3] C. Thomas, N. Balakrishnan, Performance enhancement of intrusion detection systems using advances in sensor fusion, in: Fusion’ 08: Proceedings of the 11th International Conference on Information Fusion, (2008)1671﹣1677.

[4] C. Zhou, L. Wang, Q. Zhang, X. Wei, Face recognition based on PCA and logistic regression analysis, Optik - International Journal for Light and Electron Optics,125:20(2014) 5916-5919.

[5] C.-C. Chen, L.-S. Chen, C.-C. Hsu,, and W.-R. Zeng, (2008) An information granulation based data mining approach for classifying imbalanced data, Information Sciences, 178 (16):3214–3227.

[6] F. Kuang, W. Xu, Zhang, S., A novel hybrid KPCA and SVM with GA model for intrusion detection, Applied Soft Computing 18(2014)178-184.

[7] F. Wen, J. Xiao, Z. He, X. Gong, Stock Price Prediction based on SSA and SVM, Procedia Computer Science 31(2014) 625–631

[8] G. Akritas, G. I. Malaschonok, (2004). Applications of singular-value decomposition (SVD). Mathematics and Computers in Simulation, 67, 15-31.

[9] G. Kim, S. Lee, S. Kim, A novel hybrid intrusion detection method integration anomaly detection with misuse detection, Expert Systems with Applications 41(2014)1690-1700

[10] H. Guan, J. Zhou, B. Xiao, M. Guo, T. Yang, (2013). Fast dimension reduction for document classification based on imprecise spectrum analysis. Information Sciences, 222, 147-162.

[11] J. Wang, J. Peng, Q. Liu, (2015). A classification approach for less popular webpages based on latent semantic analysis and rough set model. Expert Systems with Applications, 42(1), 642-648.

[12] K. Bhattacharyya, J. K. Kalita, ”Network anomaly detection: A machine learning perspective”, (2013),CRC Press.

[13] K. Jain, R. P. W. Duin, and J. Mao, (2000). Statistical pattern recognition: A review, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(1): 4-37.

[14] K. Uysal, and S. Gunal, (2012). A novel probabilistic feature selection method for text classification. Knowledge-Based Systems, 36, 226-235.

[15] M. Ding, Z. Tian, H. Xu, Adaptive kernel principal analysis for online featureextraction, Proc. World Acad. Sci. Eng. Technol. 59 (2009) 288–293.

[16] M. Govindarajan, R. M. Chandrasekaran, Intrusion Detection using an Ensemble of Classification Methods, Proceedings of the World Congress on Engineering and Computer Science (2012)WCECS.

[17] M. Jia, H. Xu, X. Liu, N. Wang, The optimization of the kind and parameters of kernel function in KPCA for process monitoring, Computers and Chemical Engineering 46 (2012) 94– 104

[18] M. N. Mohammed, N. Sulaiman, Intrusion Detection System Based on SVM for WLAN, Procedia Technology 1(2012)313-317

[19] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, A Detailed Analysis of the KDD CUP 99 Data Set, proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009).

[20] M.W. Berry, S. T. Dumais, and G. W. O'Brien, (1995). Using linear algebra for intelligent information retrieval. SIAM Review, 37, pp. 573-595.

[21] N. Hoque, D. K. Bhattacharyya, J. K. Kalita, “MIFS-ND: A mutual information-based feature selection method”, Expert Systems with Applications 41 (2014) 6371﹣6385.

[22] N. Hubballi, V. Suryanarayanan, False alarm minimization techniques in signature-based intrusion detection systems :A survey, Computer Communications 49(2014)1-17

[23] Nsl-kdd data set for network-based intrusion detection systems. Available on: http://nsl.cs.unb.ca/NSL-KDD/, March 2009.

[24] P. Anderson, Computer Security Threat Monitoring and Surveillance, Box 42 Fort Washington, Pa. 19034,215(1980)646-4706.

[25] S. Benferhat, A. Boudjelida, K. Tabia, H. Drias, An intrusion detection and alert correlation approach based on revising probabilistic classifiers using expert knowledge, Int. J. Appl. Intell. 38 (4) (2013) 520﹣540.

[26] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, R. Harshman, Indexing by latent semantic analysis, Information Science,(1990)2894-33

[27] S. Han, S. Cho, “Rule-based integration of multiple measure-models for effective intrusion detection Systems”, IEEE International Conference 1. 6(2003)120-125.

[28] S. Hettich, S. D. Bay, The UCI KDD Archive [http://kdd.ics.uci.edu]. Irvine, CA: University of California, Department of Information and Computer Science(1999).

[29] S. W. Lin, K. C. Ying, C. Y. Lee, Z. J. Lee, “An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection”, Applied Soft Computing, 12(10) (2012) 3285–3290.

[30] T. Liu, Z. Chen, B. Zhang, W. Ma, G. Wu, Improving text classification using local latent semantic indexing, In Proceedings of The Fourth IEEE International Conference on Data Mining, Brighton, UK, November 01 - 04, 2004

[31] V. Castelli, A. Thomasian, C.-S. Li, (2003). CSVD: Clustering and Singular Value Decomposition for approximate similarity search in high-dimensional spaces. IEEE Transaction on Knowledge and Data Engineering, 15 (3), 671-685.

[32] W. Feng, Q. Zhang, G. Uu, J. X. Huang, Mining network data for intrusion detection through combining SVMs with ant colony networks, Future Generation Computer Systems 37 (2014) 127–140.

[33] W. Lee, S. J. o. Stolf, K. W. Mok, A data mining framework for building intrusion detection models. Proceedings of the 1999 IEEE Symposium Security and Privacy 13(1999)120-132.

[34] W. Wang, R. Battiti, Identifying intrusions in computer networks with principal component analysis, in: ARES’06:Proceedings of the First International Conference on Availability Reliability and Security (2006) 270-279.

[35] W. Yang, C. Sun, L. Zhang, (2011) A multi-manifold discriminant analysis method for image feature extraction, Pattern Recognition, 44 (8), 1649-1657.

[36] X. S. Gan, J. S. Duanmu, J. F. Wang, Cong, W., Anomaly intrusion detection based on PLS feature extraction and core vector machine, Knowledge-Based Systems 40 (2013) 1–6.

[37] Z. G. Chen, H. D. Ren, X. J. Du, Minimax probability machine classifier with featureextraction by kernel PCA for intrusion detection, Wireless Communications,Netw. Mobile Comput. (2008) 1–4.