Journal of Artificial Intelligence
ISSN: 2229-3965 & E-ISSN: 2229-3973, Volume 3, Issue 3, 2012, pp.-111-116.
Available online at http://www.bioinfo.in/contents.php?id=71

APPLICATION OF ARTIFICIAL INTELLIGENCE BASED TECHNIQUES FOR INTRUSION DETECTION SYSTEMS: A REVIEW

TARAPORE N.Z.[1]*, KULKARNI D.B.[2] AND KAMAKSHI P.V.[1]
[1]Jawaharlal Nehru Technological University, Hyderabad- 500 085, AP, India.
[2]Department of Information Technology, Walchand College of Engineering, Sangli- 416 415, MS, India.
*Corresponding Author: Email- ntarapore@yahoo.com.

Received: July 12, 2012; Accepted: July 18, 2012

Abstract- Network intrusion detection is an invaluable part of an information security system. The rapid increase in the number of different types of attacks has increased the complexity involved in designing an intrusion detection system. Different techniques from various disciplines have been utilized to develop an efficient IDS. Artificial Intelligence based techniques have a major role in the development of IDS, it has several benefits over other techniques. In this paper, several AI based techniques used for the development of IDS have been reviewed. Related studies have been compared by key concept used, advantages and disadvantages of each system and the data set used.

Keywords- Artificial intelligence, Hybrid system, detection rate, intrusion detection system.

Citation: Tarapore N.Z., Kulkarni D.B. and Kamakshi P.V. (2012) Application of Artificial Intelligence Based Techniques for Intrusion Detection Systems: A Review. Journal of Artificial Intelligence, ISSN: 2229-3965 & E-ISSN: 2229-3973, Volume 3, Issue, pp.-111-116.

Introduction

Webster's dictionary defines an intrusion as the act of thrusting in, or of entering into a place or state without invitation, right or welcome. An intrusion is also defined as any set of actions that attempts to compromise the security objectives [1]. Intrusion detection is defined as the act of detecting an unauthorized activity by a computer on a network. It can be treated as a pattern recognition problem which distinguishes between network attacks and normal network behavior, taking things a step further would be the distinction between different categories of attacks.

A survey of the literature available indicates that many an intrusion detection system (IDS) has been developed implementing Artificial Intelligence (AI) based techniques. Some IDSs have been developed based on a single classification technique, while other IDSs called hybrid IDSs, have implemented more than one classification technique.

There are several objectives of this paper. The first objective is to give an introduction to the concept of IDS and give a brief taxonomy of IDSs. The second objective is to provide the review of several AI based techniques for IDS. The focus of this paper is on the core techniques in AI which include Decision Tree, Genetic Algorithm, Fuzzy Logic, Data Mining, Neural Network (NN), Support Vector Machine (SVM), Bayesian network, Markov model and several clustering techniques.

The rest of this paper is organized as follows. Section 2 discusses the taxonomy of IDS as well as the various existing IDSs. Section 3 discusses various techniques used for intrusion detection. Section 4 lists the details of AI based techniques used for IDS. Section 5 discusses hybrid approaches. Section discusses the data sets used. Section 7 gives the comparison of various AI based techniques. Section 8 concludes the paper.

Intrusion Detection System

IDS can be divided into two classes based upon data collection and storage unit:

Network based IDS: It monitors network backbones and looks for attack signatures in network traffic. It detects attacks by capturing and analyzing network traffic. The data is directly obtained from the network in the form of packets. This IDS is operating system dependent and easy to deploy on various systems.

Host based IDS: It operates on hosts and monitors the operating and file systems for signs of intrusion. It detects attacks for an individual system by using system logs and operating system audit trails. It collects the data from system calls, operating system log files, CPU utilization and application log files. The main advantage of a host based IDS is that it is operating system dependent and is very efficient to detect attacks like buffer overflow.

IDS can be divided into two classes based on the criteria of data analysis and processing unit.

Misuse Detection or Signature Based IDS: In a misuse detection based IDS, an intrusion is detected by searching for activities that correspond to known signatures of intrusions. It has a low false

positive rate, but cannot detect new types of attacks.

Anomaly Detection Based IDS: An anomaly detection based IDS detects intrusions by searching for abnormal network traffic. It can detect unknown attacks, but the false positive rate is high.

Snort, a signature based system, is the most popular network based IDS. The problem with signature based systems is that they cannot detect novel attacks. The security administrator has to update the signature database every time to keep it up to date. If there is any attack whose signature is not found in the database, then the system is unable to detect it.

Intrusion Detection Techniques

Statistical Based Techniques

Denning [2] proposes a statistical method for intrusion detection. Based upon the audit data, a profile is constructed to describe a given network user or a given task. The Gaussian models of the metrics are constructed to detect intrusions.

Knowledge Based Techniques

A Knowledge based ID technique applies the knowledge garnered regarding specific attacks and system loopholes. The IDS contains information about these loopholes and looks for attempts to exploit these loopholes. An alarm is triggered when such an attempt is detected. Any action that is not explicitly recognized as an attack is considered acceptable. The accuracy of knowledge-based IDS is considered good. Their robustness depends on the regular update of knowledge about attacks [1].

A major advantage of the knowledge-based approach is that it has the potential for very low false alarm rates. A detailed contextual analysis proposed by the IDS makes it easier for the security officer using this intrusion detection system to take preventive or corrective action.

Drawbacks include the difficulty of gathering the required information on the known attacks and keeping it up to data with new vulnerabilities and environments. A careful analysis of all vulnerabilities is required for the maintenance of the knowledge base of the intrusion detection system. It is therefore a time-consuming task.

Artificial Intelligence Based Techniques

Data Mining Techniques

Grossmann defines data mining as a field that deals with uncovering patterns, changes, associations, anomalies as well as statistically significantly structures and events in data [3]. It can also be defined as the ability to take input data and extract patterns and deviations which are not discernable to the human eye.

Lee, et al has a framework which consists of programs for learning classifiers and meta-classification, association rules for link analysis and frequent episodes for sequence analysis [4]. It also includes a support environment that enables system builders to interactively and iteratively drive the process of constructing and evaluating detection models. The learned rules replace the manually encoded intrusion patterns and profiles. The system features and measures are selected by considering the statistical patterns computed from the audit data. Deviations from rules indicate an attack on the network. Following are the major data mining techniques:

Genetic Algorithm Based Techniques

A genetic algorithm is a programming technique that mimics biological evolution as a problem solving strategy. The genetic algorithm is used to derive a set of classification rules from network audit data. The support-confidence framework is then used as a fitness function to judge the quality of each rule. The generated rules then detect or classify network intrusions in a real-time environment [5].

Song, et al. have used a hierarchical algorithm called RSS-DSS for dynamically filtering large datasets based on the concept of training pattern age and difficulty. This provides the data set of about half a million patterns in about fifteen minutes [6].

Balaji, et al. have used a genetic based intrusion detection model to learn individual user behavior and detect abnormal user activities. The User behavior is described by a 3-tuple value <Match index, Entropy index, Newness index> from commands which have actually occurred and predicted commands. The expected normal behavior 3-tuple value is compared with the calculated 3-tuple value to calculate deviations in the user behavior. The intrusion detector module computes the probability of the current command sample being intrusive from the deviations in the user behavior [7].

Chittur developed a model in which the genetic algorithm was given training data from which an empirical model of malicious computer behavior was generated. The genetic algorithm was successfully able to generate an empirical behavioral model which was successful in detecting real world testing data [8].

Fuzzy Logic Based Techniques

Luo has used fuzzy logic rules integrated with association rules and frequency episodes to classify the data [9]. A normalization step has been added to the procedure for mining fuzzy association rules developed by Kuok, et al. in order to prevent one data instance from contributing more than others [10]. A modification is also proposed to the method suggested by Mannila, et al. for mining frequency episodes to learn fuzzy frequency episodes [11].

Decision Tree Based Techniques

A decision tree is a tree with three major components, namely: nodes, arcs and leaves. Every node is labeled with a feature attribute which contains the highest information gain among attributes not yet considered in the path from the root. Each arc is labeled with a feature value for the node's feature and each leaf is labeled with a category or class. A decision tree is used to classify a data point starting at the root and traverse through the tree until a leaf node is reached. A leaf node would then provide the data classification.

Levin uses the concept of a Kernel miner, which is a data mining tool. It constructs the set of locally optimal decision trees from which it selects the optimal subset of trees used for predicting unknown cases [12]. This modeling technique reduces individual prediction results received from individual classification trees. The value of the global optimization criterion for any subset of trees is then calculated.

Quinlan developed ID3 and C4.5 which are two popular implementations of decision trees [13].

Machine Learning Techniques

Machine learning is defined as the ability of a computer program to learn and enhance the performance on a set of tasks over time [1]. These techniques focus on building a system model that enhance its performance based on previous results. Several machine learning techniques have been successfully applied to the field of intrusion detection [14].

Neural Network

A neural network is an information processing system that is inspired by the way biological nervous systems process information. A neural network comprises of a large number of highly interconnected processing elements working in unison to solve a specific problem. Each processing element is called a neuron, which is a summing element followed by an activation function. The output of each neuron is fed as the input to all the neurons in the next layer. The learning process involves finding the best set of weights for solving a problem. This is basically an optimization process [15].

Lei, et al. use the Improved Competitive Learning Network (ICLN) algorithm [16]. In this algorithm, only the winning weight vector gets its weight vector changed, following every iteration. The other weight vectors remain unchanged. Mukhopadhyay, et al. use the backpropagation neural network approach in which the error is propagated backwards, from the output layer to the hidden layer and then to the input layer [17].

Corchado, et al. uses a system called MOVCIDS (mobile visualization connectionist IDS) [18]. The system applies neural projection architectures to detect anomalous situations. Using advanced visualization features, the IDS gives an overview of network traffic. Han, et al. use a technique known as evolutionary neural network (ENN) which takes lesser time than regular neural networks since they discover the structures and weights of the network simultaneously [19].

Support Vector Machine

A Support Vector Machine (SVM) is a learning machine that plots the training vectors in high dimensional feature space. Each vector is labeled by its class. A SVM views a classification problem as a quadratic optimization problem. It avoids the “curse of dimensionality” by placing an upper bound on the margin between the different classes. This makes it easy to handle large and dynamic data sets. A SVM classifies data with the help of support vectors. These support vectors are the outline of a hyper plane in feature space, they are basically members of the set of training inputs [20].

A SVM is based on the idea of structural risk minimization which minimizes the true error over test examples. The number of free parameters used in the SVM depends on the margin that separates the data points. It does not depend on the number of input features. A SVM does not require a reduction in the number of features in order to avoid overfitting. A SVM provides a generic mechanism to fit the surface of the hyper plane to the data with the usage of a kernel function. There are different types of kernel functions which are as follows: linear, polynomial, sigmoid curve, radial basis function [20].

Li, et al. use a feature reduction method to select critical features in IDS to reduce the training time and prediction time in the IDS classifier with minimal loss of accuracy [21-23]. Sotiris, et al. use Bayesian linear model (BLM) to model the posterior class probability for test data [24]. Chang, et al. use a parallel SVM (PSVM) algorithm to reduce memory use and to parallelize both data loading and computation [25]. PSVM performs a parallel row-based Incomplete Cholesky Factorization (ICF) on the loaded data.

Zhang, et al. use a feature selection technology based on the Fisher score [26]. Sung, et al. use a technique where all the features of a data set are ranked in terms of importance [27]. The impact of a feature on the data set is determined by removing it from the data set and running the algorithm on the remaining features of the data set.

Chen, et al. reduce the number of features of the data set using Rough Set Theory prior to feeding the data to the SVM [28]. Sebald, et al. implement the concept of using the SVM to perform nonlinear equalization [29].

Bayesian Network

A Bayesian network is used to model a domain containing uncertainty. It is a directed acyclic graph (DAG) where each node represents a discrete random variable of interest. Every node comprises the states of the random variable and a conditional probability table (CPT). The CPT of a node contains the probabilities of a node being in a specific state, for a given state of its parents. In a Bayesian network, the parent-child relationship between the nodes indicates the direction of causality between the corresponding variables [30-31].

Kruegel, et al. propose an event classification scheme based on Bayesian networks [30]. This scheme improves the aggregation of different model output. It allows one to accommodate additional information.

Johansen, et al. use a Bayesian system which compares the metrics of each network traffic sample to differentiate between attacks and normal network activity [32].

Markov Model

A Markov chain is a set of states that are interconnected through certain transition probabilities which determine the topology and capabilities of the model. During the training phase, the probabilities associated with the transitions are estimated from the normal behavior of the system. The detection of anomalies is carried out by comparing the anomaly score obtained for the observed sequences with a fixed threshold.

In a hidden Markov model (HMM), the system is assumed to be a Markov process in which states and transitions are hidden. It is a double embedded stochastic process with two hierarchy levels. The states are not observable in the upper level which is a Markov process. Hoang, et al. use a simple and efficient HMM training scheme using innovative integration of multiple observations training and incremental HMM training [33].

Clustering Techniques

Clustering techniques work by grouping the observed data into clusters according to a given similarity or distance measure. The main methods for distance measurement are Euclidean distance and Mahalanobis distance. Following are the different types of

clustering:

K-means Clustering

K-means clustering is a method of cluster analysis whose aim is to partition n observations into k clusters. Every observation belongs to a cluster with the nearest mean [34].

Guan, et al. present a clustering heuristic for intrusion detection called Y-means [35]. This heuristic is based on the K-means algorithm. It overcomes a major shortcoming of K-means, which is cluster dependency.

K-NN Approach

Liao, et al. use the k-Nearest Neighbor classifier to categorize each new test case into either normal or attack [36].

Hierarchical Clustering

Hu, et al. use an unsupervised active learning framework based on hierarchical graph-theoretic clustering [37]. The dominant set clustering and spectral clustering methods are combined in a hierarchical manner.

Self Organizing Map Based Approach

A Self Organizing Map (SOM) belongs to a class of artificial neural networks which is trained using unsupervised learning. This produces a discrete, low-dimensional representation of the input space applied to the training samples, which is called a map. Self-organizing maps are different from other artificial neural networks in the sense that they use a neighborhood function to preserve the topological properties of the input space. Kayacik, et al. use the technique of a hierarchy of Self-Organizing Feature maps [38].

Jiang, et al. have proposed a novel method to compute the cluster radius threshold [39]. This method considers the outlier factor of clusters for measuring the deviation degree of a cluster. The data classification is performed by an improved nearest neighbor (INN) method. A powerful clustering method is presented for the unsupervised intrusion detection (CBUID).

Hybrid Approaches

Many researchers have suggested that the monitoring capability of the current IDS can be improved by taking a hybrid approach that consists of both anomaly as well as signature detection techniques. Hybrid techniques have been proposed by several researchers in literature.

Wang, et al. have proposed an approach which uses clustering and artificial neural networks (ANN) [40]. The fuzzy clustering technique is used to generate different training subsets. Different ANN models are used on these training subsets. A meta-learner which is a fuzzy aggregation module is used to aggregate the results. Govindrajan, et al. propose a hybrid architecture which involves ensemble and base classifiers for IDS [41]. This model uses two classification methods based on multilayer perceptron (MLP) and radial basis function (RBF).

Tsai, et al. use a technique called triangle area based nearest neighbors (TANN) in order to detect attacks [42]. k-means clustering is first used to obtain cluster centers corresponding to the attack classes. The triangle area formed by two cluster centers with a data record from the data set is then calculated and used to form a new feature signature of the data. In the last step, the k-NN classifier is used to classify similar attacks based on the new feature represented by triangle areas.

Horng, et al. use an approach which combines hierarchical clustering, a simple feature selection procedure and the SVM technique [43]. Mabu, et al. propose a technique comprising a fuzzy class association rule mining method based on genetic network programming (GNP) for detecting network intrusions [44]. GNP is an optimization technique which uses a directed graph structure instead of strings.

Hwang, et al. propose a hybrid technique which combines the advantages of low false positive rate of signature based IDS and the ability of anomaly detection system (ADS) to detect novel unknown attacks [45]. A weighted signature generation scheme is proposed to integrate SNORT with ADS. This is achieved by extracting signatures from the anomalies detected. The system extracts signatures from the output of ADS and adds them into the SNORT signature database for fast and accurate detection.

Dataset

Most of the systems mentioned in this paper have used the KDD 99 and DARPA data sets for system evaluation. The KDD 99 dataset contains 41 features which are either continuous or categorical. The KDD 99 training data set contains 22 different types of attacks. There are around five million records in the labeled dataset. Several researchers have used subsets of this database for training their systems. The testing dataset contains an additional 17 attacks, it contains 3,11,029 records.

The KDD data set suffers from inherent problems such as redundancy in the data set. Another problem is that the dataset contains an uneven distribution of attacks. There are also large contiguous portions of the data set with the same type of attack, thereby causing a bias in the training.

These problems are overcome by the NSL-KDD data set. The training set contains 1,25,973 records, the testing set contains 22,544 records [46].

Comparison of Different Techniques

A comparative study of the following techniques is presented in this section: Data mining, Decision tree, Neural network, support vector machine, clustering and hybrid. The key concept of each methodology, advantages and disadvantages of each approach, and the dataset (s) used are presented *[See Table 1]*.

Table 2- Detection rate of different methods for different classes of attacks

Method	Normal	Probe	DoS	U2R	R2L
Decision trees (Levin) [12]	99.42	84.52	97.5	11.8	7.32
Back propagation neural network (Mukkamala et al.) [47]	96.4	85.7	99.6	34.3	99.3
SVM (Mukkamala et al.) [47]	99.55	99.7	99.25	99.87	99.78
Hierarchical SOM (Kayacik et al.) [38]	95.4	95.1	64.3	10	9.9

Table 1- Comparison of Different Techniques

Technique	References (First author)	Key concept	Advantage	Disadvantage	Dataset
Data mining	Lee	Meta learning process	Reliable user anomaly detection model	Misclassification of new DOS and R2L attacks	DARPA
	Srujan	Genetic algorithm	Classification rules derived from network audit data	Low detection rate for unknown attacks	KDD 99
	Dong	Genetic programming	Technique independent of data set and structure of GP used	Low detection rate for U2R attacks	KDD 99
Decision tree	Levin	Kernel miner	Simple to use Works with several databases	Pattern finding process very time consuming	KDD 99
Neural network	Lei	Improved competitive learning neural network	Low misclassification rate, Deals with unlabeled data Classifies highly skewed data Identifies unseen patterns	Estimation of objective function value has less accuracy, convergence is slow.	Iris KDD 99 DARPA
	Mukhopadhyay	Back propagation neural network	Success rate of Level 1 is high	Training is slow Low success rate for Level 2	KDD 99
	Sang	Evolutionary neural network	Trial and error cycles not required for designing network structures	Learning time is high	DARPA
SVM	Chen	Rough set theory for preprocessing	Feature reduction	Decision table grows with data	DARPA
	Li	K-means clustering with ant colony algorithm	Training Dataset reduction	Testing not performed on KDD	KDD 99
	Zhang	Fisher score used for feature reduction	41 features reduced to 29	Unable to process entire KDD training dataset	KDD 99
	Sung	Feature ranking	Insignificant features removed	Test data set result not provided	KDD 99
Clustering	Guan	H-means+ algorithm with k-means algorithm	Overcomes some shortcomings of k-means algorithm	Selection of k for k-means algorithm	KDD 99
	Liao	k-NN method with document classifier	Implements both misuse and anomaly detection system	k-NN approach not scalable for large dataset	DARPA
	Kayacik	3 layer SOM architecture	Uses basic features of KDD only which reduces training time	Poor detection rate for R2L and U2R attacks	KDD 99
Hybrid	Horng	Hierarchical clustering and SVM	Uses complete training data with less training time	Selection of threshold value for hierarchical clustering is costly	KDD 99
	Wang	C-means clustering and ANN	High detection rate for low frequency attacks R2L, U2R	Training time high due to fuzzy clustering	KDD 99
	Mabu	Fuzzy class ARM using GNP	High detection rate	Limited extensibility	KDD 99 and DARPA

Conclusion

A comprehensive review of various AI based techniques used in intrusion detection is presented in this paper. Several studies of AI based techniques in intrusion detection systems are compared on the basis of the key concept used, their advantages and disadvantages and the data set used.

The Hybrid approach overcomes the limitations of any single approach like the anomaly or signature detection approach. The papers using the hybrid approach first reduce the data set before applying it for training. We need a data set which contains a sufficient number of instances to train the system so that the system can give a better detection rate compared to a system which is trained over a random subset of a data set.

The R2L and U2R attack classes have low detection rates since there are less number of training instances. A computationally efficient technique is required for handling large volume of data over a network.

References

[1] Kumar G., Kumar K. and Sachdeva M. (2010) *Artificial Intelligence Review*, 34, 369-387.

[2] Denning D. (1987) *IEEE Transactions on Software Engineering*, SE-13(2), 222-232.

[3] Grossman R. (1997) *Data Mining: Challenges and Opportunities for Data Mining During the Next Decade*.

[4] Lee W., Stolfo S. and Mok K. (1999) *IEEE Symposium on Security and Privacy*, 120-132.

[5] Srujan R. (2011) *International Journal of Computer Science and Security* (*IJCSS*), 5(3).

[6] Song D., Heywood M. and Zincir-Heywood A. (2005) *IEEE Transactions on Evolutionary Computation*, 9(3).

[7] Balajinath B. and Raghavan S., *Computer Communications,* 24 (12), 1202-1212.

[8] Chittur A., *High School Honors Thesis, Ossining High School*.

[9] Luo J. (199) *Master's thesis, Mississippi State University*.

[10] Kuok C., Fu A., Wong M. (1998) *SIGMOD Rec,* 27(1), 41-46.

[11] Mannila H., Toivonen H. (1996) *2nd International Conference on Knowledge Discovery and Data Mining*.

[12] Levin I. (2000) *SIGKDD Explorations*, 1(2), 67-75.

[13] Quinlan J. (1993) *C4.5: Programs for Machine Learning*.

[14] Hu W. and Maybank S. (2008) *IEEE Transactions on Systems, Man and Cybernetics*, 38(2).

[15] Beghdad R. (2008) *Computers & Security*, 27, 168-175.

[16]Lei J., Ghorbani A. (2011) *Neurocomputing*.

[17]Mukhopadhyay I., Chakraborty M., Chakrabarti S. and Chatterjee T. (2011) *International Conference on Recent Trends in Information Systems*.

[18]Corchado E. and Herrero A. (2011) *Applied Soft Computing*, 11, 2042-2056.

[19]Han S. and Cho S. (2006) *IEEE Transactions on Systems, Man and Cybernetics,* 36(3), 559-570.

[20]Mukkamala S., Janoski G. and Sung A. (2002) *Int. Joint Conf. Neural Networks*, 2, 1702-1707.

[21]Li Y., Xia J., Zhang S., Yan J., Ai X. and Dai K. (2012) *Expert Systems with Applications*, 39, 424-430.

[22]Venkatachalam V. and Selvan S. (2008) *International Journal of Simulation*, 9(1).

[23]Lee W. and Stolfo S.J. (2000) *ACM Transactions on Information and System Security*, 3(4), 227-261.

[24]Sotiris V., Tse P. and Pecht M. (2010) *IEEE Transactions on Reliability*, 59, 2.

[25]Chang E., Zhu K., Wang H. and Bai H. (2007) *Advances in Neural Information Processing Systems*, 20.

[26]Zhang X., Gu C. and Lin J. (2006) *First International Conference on Communications and Networking in China*, 1-5.

[27]Sung A., Mukkamala S. (2003) *Symposium on Applications and the Internet*.

[28]Chen R., Cheng K. and Hsieh C. (2009) *International Journal of Network Security & Its Applications*, 1(1).

[29]Sebald D. and Bucklew J. (2000) *IEEE Transactions on Signal Processing*, 48(11).

[30]Kruegel C., Mutz D., Robertson W. and Valeur F. (2003) *19th Annual Computer Security Applications Conference*, 14-23.

[31]Heckerman D. (1995) *Microsoft Research, Technical Report MSRTR*.

[32]Johansen K. and Lee S., *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.83.8479*.

[33]Hoang X. and Hu J. (2004) *12th IEEE International Conference on Networks*, 2, 470-474.

[34]MacQueen J. (1967) *5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 281-297.

[35]Guan Y., Ghorbani A. and Belacel N. (2003) *Canadian Conference on Electrical and Computer Engineering*, 2, 1983-1086.

[36]Liao Y. and Vemuri V. (2002) *Computer Security*, 439-448.

[37]Hu W., Hu W., Xie N. and Maybank S. (2009) *IEEE Transactions on Systems, Man and Cybernetics,* 39(5).

[38]Kayacik H., Zincir-Heywood A. and Heywood M. (2003) *Int. Joint Conf. Neural Networks*, 3, 1808-1813.

[39]Jiang S., Song X., Wang H., Han J. and Li Q. (2006) *Pattern Recognition Letters*, 27, 802-810.

[40]Wang G., Hao J., Ma J. and Huang L. (2010) *Expert Systems with Applications*, 37, 6225-6232.

[41]Govindrajan M. and Chandrasekaran R. *Computer Networks*, 55, 1662-1671.

[42]Tsai C. and Lin C. (2010) *Pattern Recognition*, 43, 222-229.

[43]Horng S., Su M., Chen Y., Kao T., Chen R., Lai J. and Perkasa C. (2011) *Expert Systems with Applications*, 38, 306-313.

[44]Mabu S., Chen C., Lu N., Shimada K. and Hirasawa K. (2011) *IEEE Transactions on Systems, Man and Cybernetics,* 41(1).

[45]Hwang K., Cai M., Chen Y. and Qin M. (2007) *IEEE Transactions on Dependable and Secure Computing*, 4(1).

[46]Tavallaee M., Bagheri E., Lu W. and Ghorbani A. (2009) *IEEE Symposium on Computational Intelligence in Security and Defense Applications*.

[47]Mukkamala S., Sung A., *Proceedings of* 15*th IEEE Conference on Tools.*