ISSN: xxxx-xxxx (ONLINE)
DOI:

INTRUSION DETECTION IN MANETS USING SUPPORT VECTOR MACHINE WITH ANT COLONY OPTIMISATION

M. Ramkumar, M. Manikandan, K. Sathish Kumar and R. Krishna Kumar
Department of Computer Science Engineering, Gnanamani College of Technology, India

Abstract

This paper suggested Vector Machine Service (SVM) and Intrusion Detection Ant Colony Optimization (ACO). There have been two stages of the suggested approaches. In the first level, PCA is used as a SVM preprocessor to minimize practical vector measurements and to shorten preparation time. To increasing the noise generated by interface contrast and to enhance the execution of SVM with a specific end goal. The second phase is used to distinguish identification by using the least-square support vector machine with an ACO algorithm. To order to adjust work and violence through the hunting process, ACO is using coded zooming. Ultimately, the function weights and SVM parameters are tuned concurrently in accordance with the optimal interface subset. The PCA algorithm focused on ACO with Support Vector Machine (PCA-ACO-SVM). The experiments were conducted using KDD 99 dataset which are seen as an agreed standard to assess the quality of intrusion detection to demonstrate the adequacy of the proposed method. In fact, the accurate and effective application of our suggested hybridizing approach is sensible.

Keywords:

ACO, Support Vector Machine, Principal Component Analysis, Intrusion Detection

1. INTRODUCTION

Nowadays, Internet has become an essential part in the daily activities life and it's used in various areas like banking, e-commerce, business, education, and telecommunication etc. An Intrusion describes illegal operations that breach security policy laws and then contributes to data breaches, denial of access, confidentiality and unauthorized use of services [1]. The main responsibility for the monitoring of interruptions is to defend the process structure from interloper attempting to attack a device. There are several methods for intrusion detection focused on soft computation algorithms including the vector aid (SVM), neural network (ANN), furrowed reasoning, and genetic algorithm (GA) [2].

Two major types are classified in intrusion detection methods: signature detection and abnormal detection. A signature-based or sequence that refers to an attack or risk that is recognized is the following logical misconception. SD is the way to think of finding instances to detect possible intrusions. Thanks to the use of the information obtained from specific attacks and device weaknesses. SD is otherwise referred to as the analysis of information and abuse. An anomaly-based phenomenon is the variance from a known pattern, so profiles chat over time about the normal and predicted behaviour. An exception is the lack of the observed. For certain features, accounts are also fixed and dynamic. Examples include missed sign-in attempts, device utilization, or message command. Click here for more information. Afterwards, the AD sees the normal profiles with a looked at chance to see massive attacks [3-4]. The identification technique widely used in the intrusion detection process comprises of decision-making structures, Bayesian classes, the neural networks model (NN), vector support machines (SVMs), associative ranking, K-Nearest Neighbor (KNN), inductive rules approaches. Many of these techniques upgrade the standard technique to new disruption identification data sets. The estimation of default category does not meet expectations if intrusions are significantly lower than normal behavior. In this scenario, detailed equations for such intrusion detection problems are prepared and revised by the scientists. In particular, the current computer score reliability should be enhanced so that mysterious new ambushes are remembered. When constructing competent IDS, a category remains a specific issue [5-8].

The main advantage of using the NN for device intrusion detection is that it is adaptable. False NNs are a particularly sound tool used in a few classifications, specifically for applications where, for example, design verification or identifiable evidence of a nonlinear process cannot be achieved or data is unreliable. NNs utilize vague or faulty data to identify situations that they were not ready for in the learning phase. According to the normal IDS that depend on the assault mark or a law of competence, the abovementioned kind of attacks cannot be differentiated. The other advantageous position of this technique is the signature frequency of NNs. The security of identification of property involves time-consuming attack detection. The NN frequency also allows the violation response to be fatal to any device [9-10]. The critical benefit of NN in identifying violations is that it is capable of grasping the highlights of attacks and recognizing every activity of its kind. ANN could be trained to identify negative behaviors with exceptional precision. The program would, in turn, also be able to update this training to draw attention to the instance of an attack which doesn't suit current highlights from earlier intrusions. This is an additional preference of NN, since the aggressor often seeks to steal awards from others. The Radial Basis function (RBF)-NN has difficult ties with disciplines such as interpolation, the principle of regularization, approximation of functions and noise intensity calculation. RBFNN enables the inner representation generates by the secret surface to be viewed explicitly. Multilayer perceptron's (MLPs) have significantly faster RBF NN algorithms.

2. RELATED WORKS

This paper addresses a few issues in data processing systems in the field of network security. Within this sense, many approaches and mechanisms have been developed for the limitation in conjunction with network attacks and several systems are produced to detect intrusions. The chapter addresses these structures and mechanisms briefly.

The author in [11] has suggested a more robust, less taxing RBF network for determining that layer hubs have been dissimilar, with shorter measurement period and faster pace. The author has done this by integrating 3 phases: 1. Translation of the string into numeric; 2. Deletion of redundant data; 3. Decision on the correct area for assessment. For the neural network implementation, all of these types played a decisive role. To study the principles and efficiency of the RBF network.

The author in [12] proposed to build some RBF network for the suggested ACO. It also demonstrates that the suggested strategies can generally evaluate, based on the goal function, the correct network configuration and network parameter. The ACO Algorithm was correctly designed for the continuous advancement of various requirements. ACO algorithm output and analysis have been shown to be an inspiring algorithm. This paper shows that ACO utilizes in one or another way the best integration of core advantages of all other qualified algorithms.

In this study, an advanced RBF-NN algorithm based on GA and ACO (GA-ACO-RBF) was applied in order to monitor the execution of ACO algorithm and optimization with RBF NN [13]. The paper indicated that the GA would boost RBF-NN weights and layout. ACO Architecture raised the weight and core of RBF and RBF size values. It was powered by its variable characterization.

RBF weights are encoded and the well-being of all solutions measured for the ACO target. This article has been hybridized in order to optimize the variable value. The author in [14] proposed the strategies for interference identification based on an improved genetic algorithm for the SVM kernel parameters. The basic ideas are a valid encoding, which values the model's description accuracy, the parameter enhancement for the Gaussian software SVM grouping dependent on the kernel, and the simpler parameters of the device invasion computing machine.

The author in [15] published a current and flow level analysis using SVM, and a revised technology research carried out by some scientists in this discovery field. In order to recognize a chosen, ideal component for invasion, the researcher proposed a new method. A hybrid approach incorporating channel and wrapper models for the selection of specific apps is the path forward. This lowered dataset draws on the execution and detection of the SVM Based Detection Method [16].

The optimal extraction method for components, referred to as OA PCA, was explored as well as the possibility to classify multi-classification epileptic EEG signals with a suitable category classifier. There are many benefits in this article, such as high order implementation for all classifiers attempted and very small FAR.

In [17], the author reviewed the identification of different kinds of IDS assaults was a complicated problem which involved a collection of vast IDS knowledge datasets. Delegated research from a comprehensive information array plays a key role in detecting intrusion in the field of network security. The suggested IDS method uses testing, which we list as the perfect distribution least square vector aid system. A KDD 99 dataset is examined and accepted for evaluating the proposed system, which is evaluated as a review benchmark for any IDS solution. The test shows that the proposed techniques were necessary to discriminate between the three specific phases of the proposed IDS display [18].

To order to get to grips with the data package, the representative weighted credits are pre-processed and the context algorithm is implemented. At that point, IG is used for highlight choice and SVM is used for order. The swarm intelligence algorithm (PSO) or the Artificial Bee Colony (ABC) is used to pick parameters for SVM.

In the last several years, while various intrusion detection techniques focused on SVM have been introduced, the algorithms above still encounter the harmful effects of certain inadequate components. Conventional element description (e.g. PCA) fails to take into account different touchy aspects, resulting in an optimal category without effect. When GA is used to develop the intrusion detection system focused on SVM, the preparation time is longer; the error rate is higher, while the optimal element subset is chosen. The value of the highlights is not organized with the selection of the ideal element section.

This paper introduced a novel approach to the identification of low visit attacks and the accuracy of the detection, combining ACO-SVM and PCA. This methodology allows PCA to map the characteristics of the high dimension in the output space into another lower-specific measurement area and concentrates the core highlights of the uniform data. With the ultimate goal of shortening training times and improving the performance of SVM classification displays, an optimized RBF feature is generated for respect to the Gaussian component and ACO is used to simplify SVM parameters.

3. PROPOSED METHOD

The Ant Colony Optimization (ACO) algorithm has a mixture of decentralized estimation, autocatalysis (positive feedback) and competitive ambition to find the ideal solution to problems in tandem with optimisation. This method aims to emulate the actions of the ant in the real world. Since its introduction, the ACO algorithm has received much attention and has been incorporated in many optimization problems, namely the network routing, traveling salesman, quadratic assignment, and resource allocation problems.

A general outline of the ACO algorithm is given below.

```
Algorithm ACO meta heuristic();
    while (termination criterion not satisfied)
        ant generation and activity();
        pheromone evaporation();
        daemon actions(); “optional”
    end while
end Algorithm
```

The ants in ACO algorithm have the following properties:

Step 1: Each ant searches for a minimum cost feasible partial solution.

Step 2: An ant k has a memory M^k that it can use to store information on the path it followed so far. The stored information can be used to build feasible solutions, evaluate solutions and retrace the path backward.

Step 3: An ant k can be assigned a start state s^k_s and more than one termination conditions e^k.

Step 4: Ants start from a start state and move to feasible neighbor states, building the solution in an incremental way. The procedure stops when at least one termination condition e^k for ant k is satisfied.

Step 5: An ant k located in node i can move to node j chosen in a feasible neighborhood N^k_i through probabilistic decision rules. This can be formulated as follows:

Step 6: An ant k in state $sr =< s_{r-1}; i >$ can move to any node j in its feasible neighborhood N^k_i , defined as $N^k_i = \{j \mid (j \in Ni) \Lambda (< sr, j > \in S)\}$ $sr \in S$, with S is a set of all states.

Step 7: A probabilistic rule is a function of the following.

Step 8: The values stored in a node local data structure $Ai = [a_{ij}]$ called *ant routing table* obtained from pheromone trails and heuristic values,

Step 9: The ant's own memory from previous iteration, and

Step 10: The problem constraints.

Step 11: When moving from node i to neighbor node j, the ant can update the pheromone trails τ_{ij} on the edge (i, j).

Step 12: Once it has built a solution, an ant can retrace the same path backward, update the pheromone trails and die.

4. OPTIMIZING THE SVM MODEL PARAMETERS WITH ACO

ACO is a follow-up tool used to search fuzzy solutions to changes or research issues in the manufacturing system. ACO algorithm, a global search heuristic, are a specific class of evolutionary algorithm that utilization strategies propelled by evolutionary science, for example, legacy, transformation, determination. The increased role of ACO in a wide range of applications has been recognized. The proposed SVM with ACO algorithm is shown in Fig.1.

Table.1. Five training and testing data

Datasets	Training Set		
	Normal (%)	Abnormal (%)	Total
DS1	84	16	12599
DS2	90	10	11010
DS3	55	45	8998
DS4	94	6	10658
DS5	78	22	6820
Datasets	**Test Set**		
	Normal (%)	**Abnormal (%)**	**Total**
DS1	73	27	11052
DS2	35	65	11428
DS3	58	42	12833
DS4	85	15	11578
DS5	65	35	12305

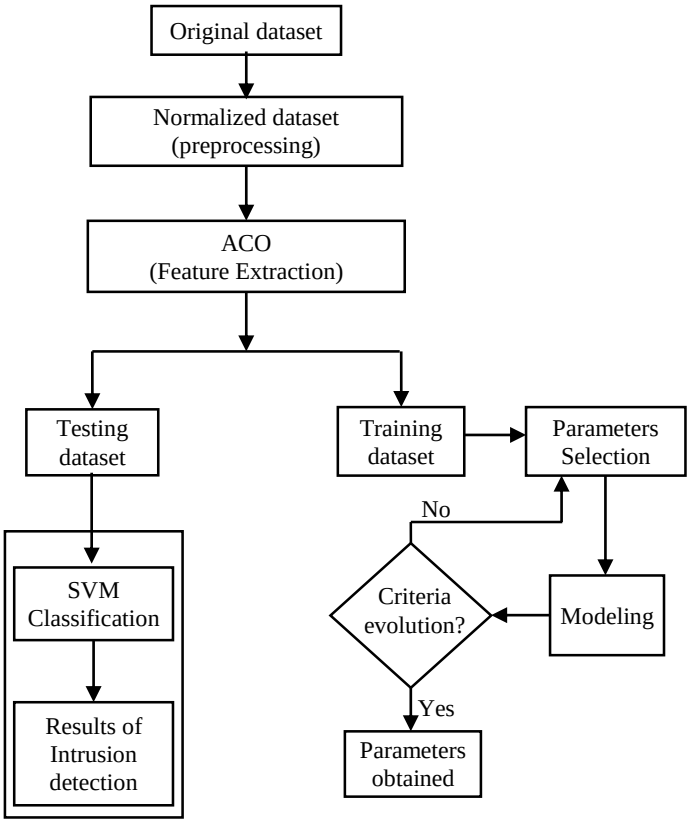

Fig.1. Proposed SVM model with ACO Algorithm

This paper uses ACO algorithm are used to enhance the parameters, σ, ε and C of SVM. Mean Absolute Percentage Error (MAPE) is utilized for evaluating fitness [27]–[29]:

$$\text{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left[\frac{a_i-d_i}{a_i}\right]\times 100\% \quad (1)$$

where, a_i and d_i correspond to the actual and estimated values, respectively. N is the number of classification period. ACO is used to generate a smaller MAPE by looking for a better three-parameter combination in SVM. Figure explains the technology to advance the SVM parameters.

In this sector, we have selected the KDD 99 dataset to form the study and exercise array. Table 1 contained five sets of data. The identification is provided by the SVM in MATLAB in this article. We analyze an area of work that is performed under the same terms, with Intel(R) Core(TM) i5-2600. We offer our tests the precision, recall and F-value from [30] and not only the reliability of the provided data set, anything else but the toughness of deliberately choosing a sample size to achieve high accuracy. The precision, reminder and F-value are described as:

$$Detection\ Rate = \frac{TP}{TP+FP} \quad (2)$$

$$False\ Alarm\ Rate = \frac{TP}{TP\ +\ FN} \quad (3)$$

$$CC = \frac{TP\times TN - FP\times FN}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(TN+FN)}} \quad (4)$$

Some of the Intrusion monitoring performance indicators are TP, FP, TN, and FN where TP is reliable in forecasting normal

behavior, FP is indicating that unusual behavior is considered normal, FN means normal behavior has not been found to be irregular and TN is successful in detecting abnormal behavior.

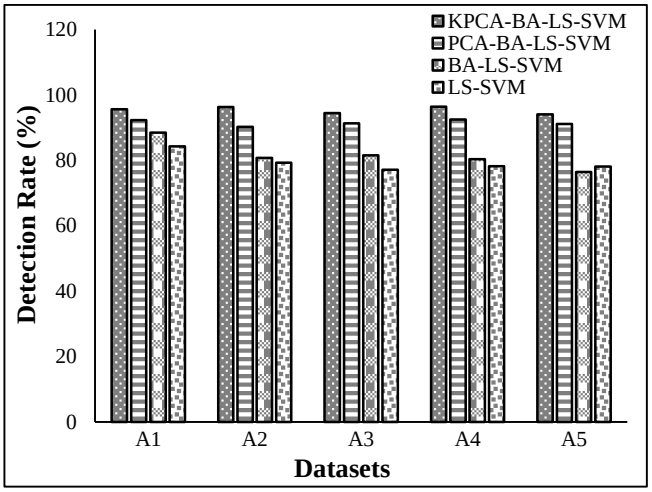

Fig.2. Detection Rate

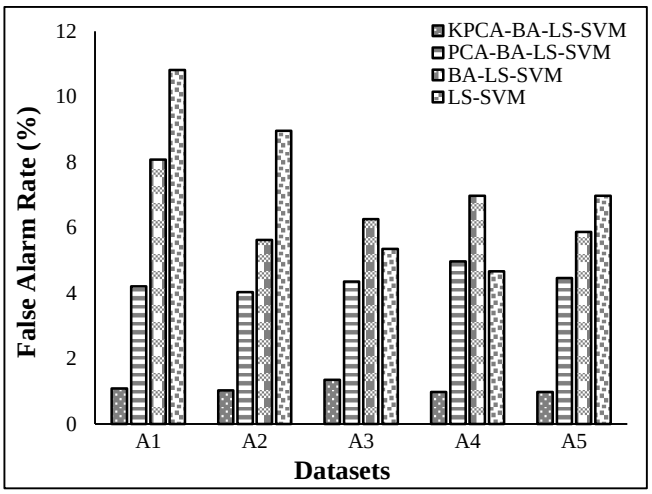

Fig.3. False Alarm Rate

The supporting experiments were carried out to validate the feasibility of the current PCA-ACO-SVM. Just outside the ACO, the group that we had has been arbitrarily split into two subsets; each subset contains information from both normal and abnormal classes: one as the activity collection and the other as the study set.

In contrast, 5 data sets, called from A1 through A5, are selected as training sets by randomly. Second, the standard and attack records are selected from the test sub-set to be generated with a comparable number.

Related experiments to verify the feasibility of the PCA-ACO-SVM were performed. In this context, just outside the ACO, the group we have in Table.1 was arbitrarily split into two subsets, each sub-set included the data of both the normal class as well as the unusual one as a training set. In fact, pick 5 datasets as a training set randomly from the testing group designated A1 to A5. Thirdly, the standard and attack logs with a similar number are chosen from the sample subset to shape the test array. The results are given in Fig.2-Fig.6.

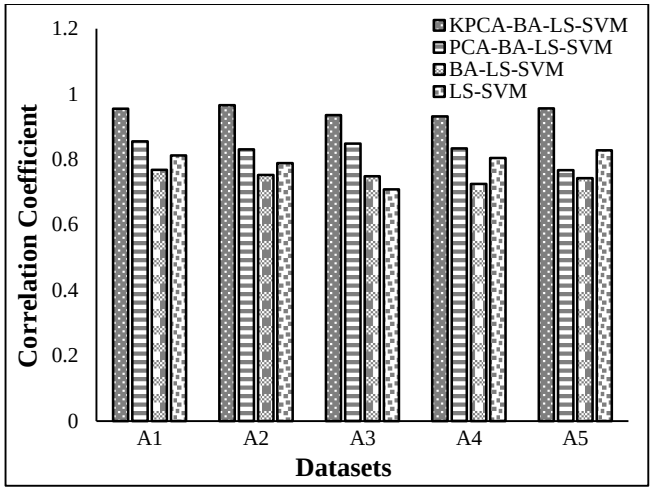

Fig.4. Correlation Coefficient (CC)

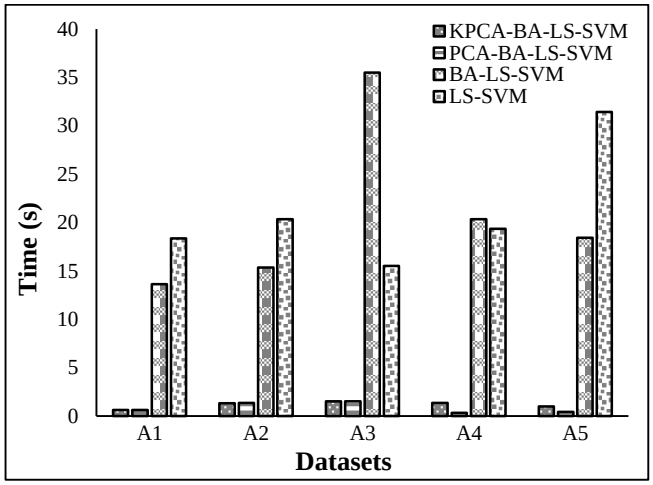

Fig.5. Total Training Time

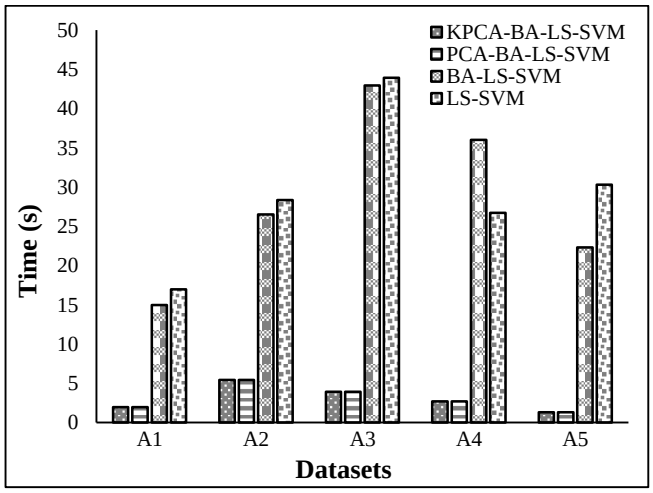

Fig.6. Total Testing Time

5. CONCLUSIONS

This paper proposes to identify intrusions by introducing a novel hybrid PCA-SVM with ACO design. PCA-ACO-SVM demonstrates that the essential aspects of intrusion detection

software are isolated from PCA, and that multi-layer SVM identification is used to determine whether the behavior is an attack. Because of the Gaussian kernel function kernel is generated to abbreviate the preparation time and to progress the SVM classification execution, ACO is used to pick correct parameters for SVM classifiers, which do not override or override the SVM proof that occurs because of the unknown parameters to ensure the SVM classification. We will build more algorithms for further research in conjunction with kernel systems with several other identification methods for expected analysis and on-line intrusion detection.

REFERENCES

[1] J. Jabez and B. Muthukumar, "Intrusion Detection System (IDS): Anomaly Detection using Outlier Detection Approach", *Procedia Computer Science*, Vol. 48, pp. 38-346, 2015.

[2] Gulshan Kumar, Krishan Kumar and Monika Sachdeva, "The Use of Artificial Intelligence Based Techniques for Intrusion Detection: A Review", *Artificial Intelligence Review*, Vol. 34, No. 4, pp. 369-387, 2010.

[3] Zahra Bazrafshan, Hashem Hashemi, Seyed Mehdi Hazrati Fard and Ali Hamzeh, "A Survey on Heuristic Malware Detection Techniques", *Proceedings of 5th International Conference on Information and Knowledge Technology*, pp. 113-120, 2013.

[4] N. Ye, S.M. Emran, Q. Chen and S. Vilbert, "Multivariate Statistical Analysis of Audit Trials for Host-Based Intrusion Detection", *IEEE Transactions on Computers*, Vol. 51, No. 7, pp. 810-820, 2002.

[5] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia Fernandez and E. Vazquez, "Anomaly based Network Intrusion Detection: Techniques, Systems and Challenges", *Computer and Security*, Vol. 28, pp. 18-28, 2009.

[6] C. Kruegel, D. Mutz, W. Robertson and F. Valeur, "Bayesian Event Classification for Intrusion Detection", *Proceedings of International Conference on Annual Computer Security Applications*, pp. 14-23, 2003.

[7] D.Y. Yeung and Y. Ding, "Host-Based Intrusion Detection using Dynamic and Static Behavioral Models", *Pattern Recognition*, Vol. 36, No. 1, pp. 229-243, 2003.

[8] A.M. Cansian, E. Moreira, A. Carvalho and J.M. Bonifacio, "Network Intrusion Detection using Neural Networks", *Proceedings of International Conference on Computational Intelligence and Multimedia Applications*, pp. 276-280, 1997.

[9] T.R. Srinivasan, R. Shanmugalakshmi and B. Madhusudhanan, "Dynamic Remote Host Classification in Grid Computing using Clonalg", *Proceedings of International Conference and Workshop on Emerging Trends in Technology*, pp. 198-201, 2010.

[10] K.S. Devikrishna and B.B. Ramakrishna, "An Artificial Neural Network Based Intrusion Detection System and Classification of Attacks", *International Journal of Engineering Research and Applications*, Vol. 3, No. 4, pp. 1959-1964, 2013.

[11] Jing Bi, Kun Zhang and Xiaojing Cheng, "Intrusion Detection Based on RBF Neural Network", *Proceedings of International* Symposium *on Information Engineering and Electronic Commerce*, pp. 357-360, 2009.

[12] S. Velliangiri and J. Premalatha, "Intrusion Detection of Distributed Denial of Service Attack in Cloud", *Cluster Computing*, Vol. 25, No. 2, pp. 1-9, 2017.

[13] S. Velliangiri and R. Selvam, "Investigation Distributed Denial of Service Attack Classification Using MLPNN-BP and MLPNN-LM", *Journal of Computational and Theoretical Nanoscience*, Vol. 15, No. 9-10, pp. 2764-2768, 2018.

[14] S. Velliangiri, R. Cristin and P. Karthikeyan, "Genetic Gray Wolf Improvement for Distributed Denial of Service Attacks in the Cloud", *Journal of Computational and Theoretical Nanoscience*, Vol. 15, No. 7-8, pp. 2330-2335, 2018.

[15] S. Siuly and Y. Li, "Designing A Robust Feature Extraction Method Based on Optimum Allocation and Principal Component Analysis for epileptic EEG Signal Classification", *Computer Method and Programs in Biomedicine*, Vol. 119, No. 1, pp. 29-42, 2015.

[16] M.D. Enamul Kabir and Jiankum Hu, "A Statistical Framework for Intrusion Detection System", *Proceedings of 11th International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 941-946, 2014.

[17] A.C. Enache and V.V. Patriciu, "Intrusions Detection Based On Support Vector Machine Optimized with Swarm Intelligence", *Proceedings of 9th IEEE International Symposium on Applied Computational Intelligence and Informatics*, pp. 153-158, 2014.

[18] I.T. Jolliffe, "*Principle Component Analysis*", Springer, 1986.

[19] Z.G. Chen, H.D. Ren and X.J. Du, "Minimax Probability Machine Classifier with Feature Extraction by Kernel PCA for Intrusion Detection", *Wireless Personal Communications*, Vol. 10, No. 3, pp. 1-7, 2008.

[20] M. Ding, Z. Tian and H. Xu, "Adaptive Kernel Principal Analysis for Online Feature Extraction", *World Academy of Science, Engineering and Technology*, Vol. 59, pp. 288-293, 2009.

[21] V. Vapnik, "*The Nature of Statistical Learning Theory*", Springer, 1995.

[22] J.A.K. Suykens, T.V. Gestel, J.D. Brabanter, B.D. Moor and J. Vandewalle, "*Least Square Support Vector Machine*", World Scientific Press, 2002.

[23] S. Mandal, G. Saha and R.K. Pal, "Recurrent Neural Network Based Modeling of Gene Regulatory Network using ACO Algorithm", *Journal of Bioinformatics and Computational Biology*, Vol. 15, No. 4, pp. 1-12, 2017.

[24] X.S. Yang and X. He, "ACO Algorithm: Literature Review and Applications", *International Journal of Bio-Inspired Computation*, Vol. 5, No. 3, pp. 141-149, 2013.

[25] X. Meng, X.Z. Gao and Y. Liu, "A Novel Hybrid ACO Algorithm with Differential Evolution Strategy for Constrained Optimization", *International Journal of Hybrid Information Technology*, Vol. 8, No. 1, pp. 383-396, 2015.

[26] D. Srivastava, L. Bhambhu, "Data Classification using Support Vector Machine", *Journal of Theoretical and Applied Information Technology*, Vol. 12, No. 1, pp. 1-7, 2010.

[27] C.W. Hsu, C.C. Chang and C.J. Lin, "A Practical Guide to Support Vector Classification", Available at: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

[28] R. Chen, D.Y. Sun, D.T. Qin, F.B. Hu, "A Novel Engine Identification Model Based on Support Vector Machine and Analysis of precision-Influencing Factors", *Journal of Central South University of Technology*, Vol. 41, No. 4, pp. 1391-1397, 2010.

[29] K.K. Gupta, B. Nath and R. Kotagiri, "Layered Approach using Conditional Random Fields for Intrusion Detection", *IEEE Transactions on Dependable and Secure Computing*, Vol. 7, No, 1, pp. 35-49, 2010.

[30] S.X. Wu and W. Banzhaf, "The Use of Computational Intelligence in Intrusion Detection Systems: A Review", *Applied Soft Computing*, Vol. 10, No. 1, pp. 1-35, 2010.