ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

IMPLEMENTATION OF SEQUENTIAL SVM CLASSIFIER TO IMPROVE RESPONSE TIME AND DETECTION RATE

Pooja Champaneria*, Prof. Bhavin Shah, Asst. Prof. Krunal Panchal
*Student of Final Year M.E.(C.E.), L.J.Institte of Engineering and Technology, Ahmedabad.
M.C.A. Programme, L.J. Institute of Management Studies, Ahmedabad.
Department of Computer Engineering, L.J. Institute of Engineering & Technology, Ahmedabad, India.

ABSTRACT

It is important to protect the assets we exchange on network that may be harmed by malicious activities. For detection of malicious activities we use Intrusion Detection System (IDS). Focus of this paper is on IDS using SVM. SVM is used as it has property of high scalability and high speed of classification which proves it efficient for IDS. The survey result shows problem of low detection rate and high response time when using traditional SVM. The problem of surveyed literature is overcome by implementing SVM-SMO model proposed in this paper and using appropriate pre-processing method. The model optimizes the Lagrange multiplier and finds support vectors which SVM algorithm uses for classification. The SVM-SMO model in this paper is implemented for sequential as well as parallel approach. The weight updating module is used by Sequential approach which prioritizes SVM classifier for improved performance. The experimental result shows improvement of 3.94% and 1.85s in detection rate and response time by implementing SVM-SMO model in sequential approach as well as improvement of 8.91s in response time using parallel approach.

KEYWORDS Classification, Intrusion Detection System(IDS), Multiclass SVM (MCSVM), Neural Network(NN), Sequential Minimal Optimization (SMO), Support Vector Machine(SVM).

INTRODUCTION

Due to rapid growth of information exchange and electronic commerce in the recent decade, we use Internet to carry out most of our transaction activities. We share our private, public confidential and business related assets using Internet connection. So, prime importance is protection of these types of important assets against intrusive activities. To prevent the intrusive behavior of network attacks, detection of it is of prime importance. Various open source IDS can be used to detect anomalies as mentioned in paper [13] for protecting important assets. Intrusion Detection mechanism can be defined as the process of monitoring the events occurring in a computer system or network and analyzing the same for the signs of intrusion [1]. Intrusion detection system is classified mainly as Host based and Network based Intrusion Detection System [5]. Technique for IDS is classified as misuse based detection and anomaly detection [5].

From the literatures surveyed for various neural network techniques, it was found that SVM has property of high scalability and high speed [15] of classification which proves it efficient for IDS. So focus is laid on Support Vector Machine for incorporating security mechanism of the proposed system in this paper.

RELATED WORK

This paper discusses implementation of IDS using SVM for KDDCUP'99 dataset. The main reason behind implementing IDS using SVM system in the network is to increase the detection rate of attacks and minimize the generation of false alarm. For experimental purpose KDDCUP'99 dataset is used as it makes the task of result comparison easy**.**

From literatures reviewed it was found that feature reduction plays important role in evaluating performance. In paper [2], PCA technique for feature reduction is implemented. Various feature subsets are formed and the

experiment shows that by applying feature reduction technique, improvement in detection rate is observed. In paper [3], a fuzzy membership function is assigned to the data that are supplied for training. By applying fuzzy factor to each training data during pre-processing, improvement in performance is observed. Also it is observed that more time is required to train 41 features. So we can use proper feature selection technique which improves the performance of the system proposed in this paper.

Motivation behind the proposed system in this paper was found after carrying out the literature survey in [14]. Major challenges found in literature reviewed are: response time [4][5], detection rate [3][4][5][6][7] and false alarm [2][5]. This paper focuses on improving detection rate and response time using SVM-SMO model by implementing SVM classifier using sequential and parallel approach.

SUPPORT VECTOR MACHINE

SVM algorithm was invented by Vladimir N. Vapnik and the current standard incarnation (soft margin) was proposed by Corinna Cortes and Vapnik in 1993 and published in 1995 [16]. Support vector machine is a kind of statistical learning theory based on the principle of structural risk minimization [8][9]. Support vector machines, one of the techniques of neural network are a set of supervised learning methods used for classification, regression and outlier's detection [24]. SVM aims at separating the classes using optimal boundary and the optimal boundary selected should be such that it maximizes the distance between two classes [17] as shown in Figure 1. SVM is a binary classifier but in the problems where data needs to be classified into more than two classes multiclass SVM is used. In case of linear dataset, a straight line or a hyper plane can be used for data classification where as in case of non-linear dataset, the data in SVM is classified on the basis of the kernel function used. Various kernel functions such as linear kernel, Polynomial kernel, RBF(Radial Basis Function) kernel and Sigmoid kernel can be used for dataset classification.

Binary SVM

SVM is a binary classifier which classifies the data into two classes i.e either positive class or negative class [15].

Multiclass SVM

To overcome the limited scope of binary SVM classifier, multi-class SVM classifier is adopted which performs classification using the following approach: a) one versus one approach b) one versus rest approach c) Direct acyclic graph based approach [18].

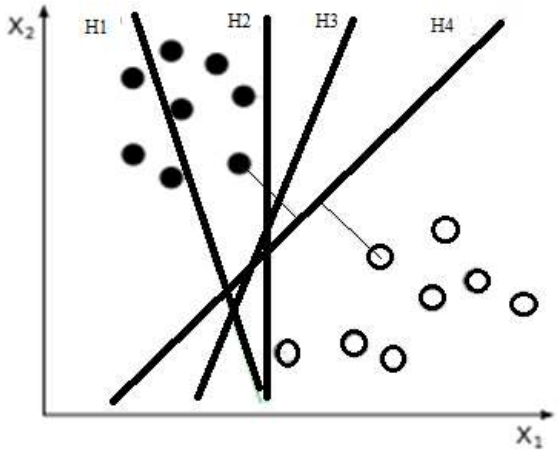

Figure 1: H:boundary.H1 does not separate the classes. H2 does, but only with a small margin. H4 separates them with the maximum margin [20][22][23]

Learning of SVM

From the literatures surveyed [14][16][18], learning process of SVM can be described in following steps:

1) Initialize the input vectors.
2) If input vectors are linearly separable, Goto 4; Else Goto 3.
3) Transform input vectors into high dimensional feature space that is hidden from both input and output.

4) Select optimal decision function that separates input vectors into target class.
5) Computation of kernel function which classifies data into target class.

Repeat above steps until all input vectors are classified.

SEQUENTIAL MINIMAL OPTIMIZATION

Sequential Minimal Optimization (SMO) is a simple algorithm that can quickly solve the SVM QP problem without any extra matrix storage and without using numerical QP optimization steps at all[19]. SMO decomposes the overall QP problem into QP sub-problems, to ensure convergence. The noteworthy features of the SMO algorithm are computational speed and ease of implementation.

PROPOSED MODEL

The proposed model for sequential and parallel approaches is shown in Figure 2 and 3 respectively. For training and testing of the proposed model, we have used KDDCUP'99 dataset, as it is benchmark for the IDS system. The main reason behind using KDDCUP'99 dataset is they are standardized dataset and the experimental results can be compared with other IDS approaches. Other reason is it consists of 41 features and wide variety of attacks which is difficult to find in any other dataset. Using KDDCUP'99 dataset we have 41 input and 1 output [11]. As KDDCUP'99 dataset contains symbolic data so we need to pre-process the data set before using it as input for the proposed model. In the proposed system we are working with Normal, DOS, Probe, U2R and R2L attacks [25].

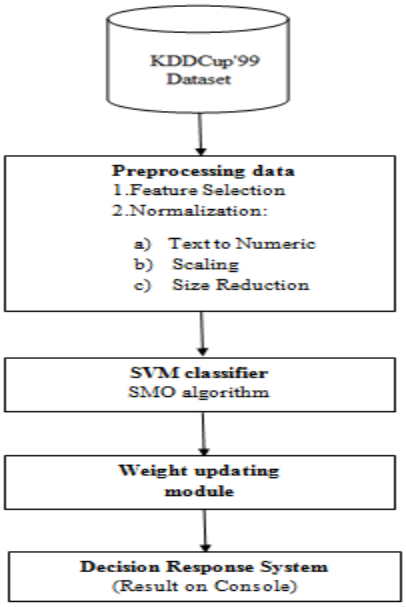

Figure 2: Proposed Model for sequential approach

In order to improve the performance, preprocessing of dataset is an essential step. The steps used for preprocessing dataset as mentioned in the model are categorized as follow: i) Feature selection and ii) Normalization. For feature selection, from the survey of the paper [10][12], 22 features will be used from the available 41 features of KDDCUP'99 dataset to carry out the experiments. Various methods are available for feature selection process which uses approximately 6 to 10 features for experimental purpose. As KDDCUP'99 is standardized dataset, reducing the features from 41 to a small count of 6 to 10 results in degrading the performance. So from the literature survey carried out for feature selection, 22 features are selected to carry out experiment for the proposed model in this paper. From paper [21] various operations like text to numeric conversion, scaling and size reduction is carried out to normalize the dataset. Text to numeric conversion converts the string value into numeric form where as scaling scales the large values in dataset so that it lies in the range of [-1,1] and size reduction removes the repeated values from the dataset.

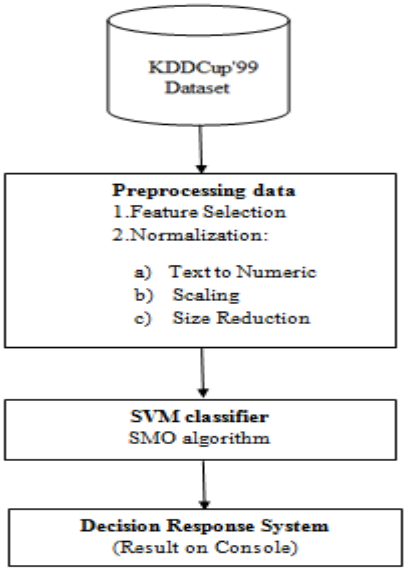

Figure 3: Proposed Model for parallel approach

In order to improve the performance, preprocessing of dataset is an essential step. The steps used for preprocessing dataset as mentioned in the model are categorized as follow: i) Feature selection and ii) Normalization. For feature selection, from the survey of the paper [10][12], 22 features will be used from the available 41 features of KDDCUP'99 dataset to carry out the experiments. Various methods are available for feature selection process which uses approximately 6 to 10 features for experimental purpose. As KDDCUP'99 is standardized dataset, reducing the features from 41 to a small count of 6 to 10 results in degrading the performance. So from the literature survey carried out for feature selection, 22 features are selected to carry out experiment for the proposed model in this paper. From paper [21] various operations like text to numeric conversion, scaling and size reduction is carried out to normalize the dataset. Text to numeric conversion converts the string value into numeric form where as scaling scales the large values in dataset so that it lies in the range of [-1,1] and size reduction removes the repeated values from the dataset.

After pre-processing we are using SMO algorithm to find the Support Vector and the optimized value of Lagrange multiplier. Lagrange multiplier helps selecting the optimal decision boundary allowing the minimization of error rate for classifying the data. These values of SMO algorithm are used while implementing SVM classifiers. As SVM is a binary classifier, to classify our dataset into five classes we need to implement multi-class SVM. One versus all approach is used for multi class classification for the proposed SVM-SMO model. Classification of the data in SVM depends on:

$$f(x) = sign\left(\sum_{i=1}^{l} \alpha_i \, y_i k(x, x_i) + b\right) \quad (1)$$

where α_i is the Langrange multiplier, b is the bias, y denotes the class label, k denotes the kernel function that will be used and x denotes the input data. After SVM algorithm classifies the data using sequential approach of the proposed model, weight updating module is implemented which priorities the SVM classifier. By implementing weight module for sequential approach improvement in performance is observed. Lastly result is displayed on the screen after performing classification.

The performance measure is done on the basis of True Positive, True Negative, False Positive, False Negative, Accuracy, Precision and Detection rate of individual attacks. Following are the equation used for calculating performance parameters:

True Positive (TP):

$$\text{TP} = \frac{\text{Correct Detected Attacks}}{\text{Total no.of Attacks}} \quad (2)$$

False Positive (FP):

$$FP = \frac{\text{No.of Normal Detected as Attack}}{\text{Total no.of Normal}} \quad (3)$$

True Negative (TN):

$$TN = \frac{\text{Correct Detected Normal}}{\text{Total no.of Normal}} \quad (4)$$

False Negative (FN):

$$FN = \frac{\text{No.of Intrusion Detected as Normal}}{\text{Total no.of Attacks}} \quad (5)$$

Accuracy:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

Detection Rate:

$$DR = \frac{\text{correctly detected attacks}}{\text{Total number of attacks}} \quad (8)$$

EXPERIMENTAL RESULTS

We have carried out two experiments for the proposed SVM-SMO model. In first experiment, performance is evaluated for various features for all kernel functions. In second experiment, comparison of sequential and parallel approach for the proposed model is evaluated.

Experiment-1 Results:

The results in Figure 4 for accuracy and response time implemented using RBF, Polynomial and Sigmoid kernel is shown for 41 as well as 22 features for sequential approach. The result shows that by using 22 features for sequential approach, we can observe improvement in accuracy and response time compared to 41 features. Also the performance of RBF kernel is best among all the three kernels.

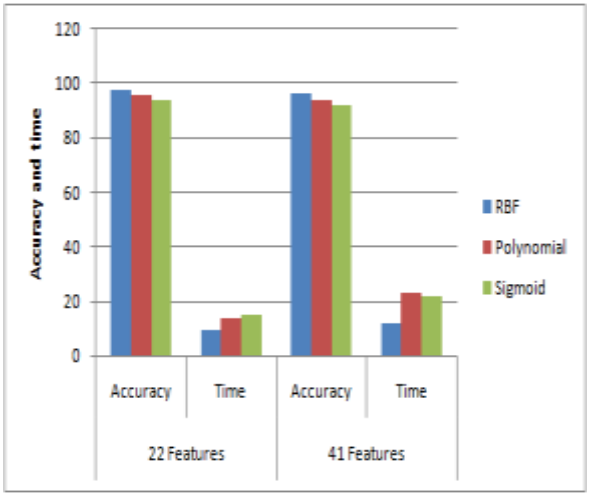

Figure 4: Accuracy and time of all kernel for 41 and 22 features using sequential approach

Experiment-2 Results:

The experimental results in Figure 5 and Table 1 show response time by implementing RBF, Polynomial and Sigmoid kernel for sequential as well as parallel approach. Results show that by implementing kernel function in parallel, we can achieve improvement in response time.

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

Table 1: Response time of all kernel for sequential and parallel execution

	Kernels		
	RBF	**Polynomial**	**Sigmoid**
Linear	9.024s	9.599s	22.462s
Parallel	1.951s	2.139s	4.604s

Comparing sequential and parallel approach, we can observe that we get better response time while implementing parallel approach. But parallel approach cannot be adopted in every scenarios of real life application due to scarcity of resources like disk space, storage memory, etc. So to avoid this problem sequential approach is used in most real life application which implements the weight updating module for improving the performance. The comparison of sequential approach with model [3] shows that performance of SVM-SMO model is better than performance using model [3] shown in Table 2 which proves that sequential approach of this paper gives better performance.

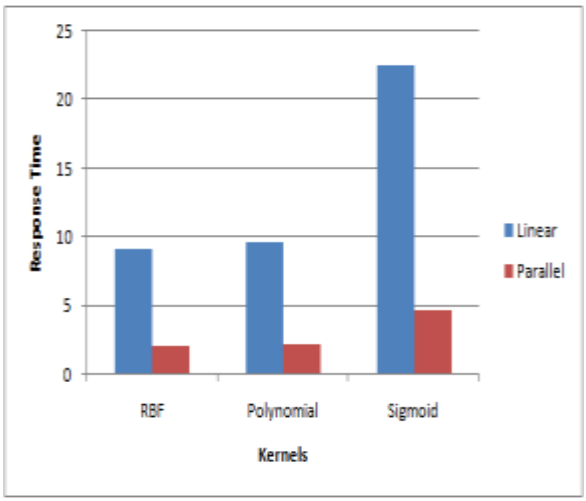

Figure 5: Response time for sequential and parallel execution of all kernels

COMPARISON ANALYSIS

The existing system for MSVM [3] and FMSVM [3] classifier is compared with proposed system of the paper using SVM-SMO algorithm shown in Figure 6 and Table 2. Improvement in detection rate and response time is observed while implementing SVM-SMO model using sequential approach.

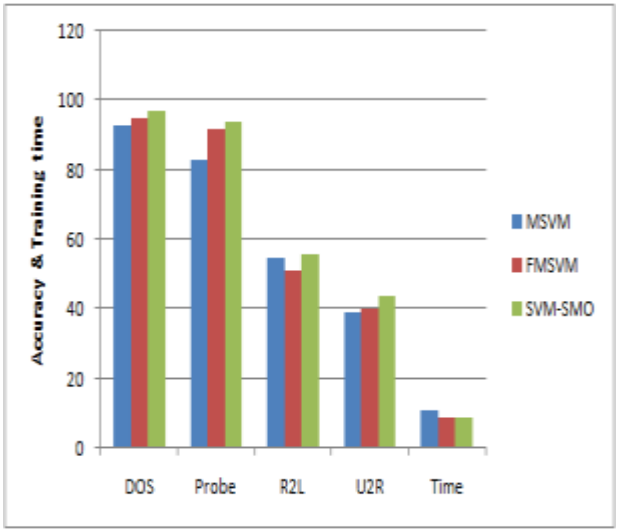

Figure 6: Comparison of existing and proposed system in this paper

Table 2: Comparison with existing system

Classifier	DOS (%)	Probe (%)	R2L (%)	U2R (%)	Time sec
MSVM [3]	93.00	83.00	55.00	39.00	10.87
FMSVM[3]	95.00	92.00	51.00	40.00	9.11
SVM-SMO	96.94	93.97	55.69	43.73	9.02

CONCLUSION

To detect anomalies in network data, SVM-SMO model proposed in this paper is used after pre-processing the dataset. Properties of SVM like high scalability and high speed of classification proves it efficient for IDS. SMO algorithm optimizes the value of Lagrange multiplier and finds Support Vectors which SVM algorithm further uses for classification of network data. Also, selection of kernel function plays important role for classifying the network data accurately. The paper shows evaluation of SVM-SMO model for sequential and parallel approach. The experimental result shows that after pre-processing the dataset and proper selection of kernel function, improvement in performance is observed. For sequential approach of SVM-SMO model, improvement of 3.94% in detection rate and 1.85s in response time is observed where as improvement of 8.91s in response time is observed for parallel approach.

DECLARATION

All the content of the current paper is purely written by Pooja Champaneria under guidance of Bhavin Shan and Krunal Panchal. The content of this paper is written by Author1 (Pooja Champaneria) while Author2 (Prof. Bhavin Shah) had guided the work and Author3 (Asst. Prof. Krunal Panchal) had reviewed this paper. Hence Author1 is responsible for the content and issues related with plagiarism.

REFERENCES

[1] D.S Bauer, M.E Koblentz "NIDX- an expert system for real-time network intrusion detection",IEEE, Proceedings of the Computer Networking Symposium, 1988. pp. 98-106, ISBN:0-8186-0835-8.

[2] Noreen Kausar, Brahim Belhaouari Samir, Suziah Sulaiman, Iftikhar Ahmad, Muhammad Hussain "An Approach towards Intrusion Detection using PCA Feature Subsets and SVM"International Conference on Computer & Information Science (ICCIS), IEEE 2012,ISBN:978-1-4673-1937-9.

[3] Lei Li, Zhi-ping Gao, Wen-yan Ding ,"Fuzzy multi-class support vector machine based on binary tree in network intrusion detection", International Conference on Electrical and Control Engineering (ICECE), IEEE 2010,ISBN:978-1-4244-6880-5.

[4] Guan Xiaoqing, Guo Hebin, Chen Luyi, "Network intrusion detection method based on Agent and SVM", The 2nd International Conference on Information Management and Engineering (ICIME), IEEE 2010, ISBN: 978-1-4244-5263-7.
[5] Singh S, Singh J P, Shrivastva G "A hybrid artificial immune system for IDS based on SVM and belief function", Fourth Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT),IEEE 2013, ISBN: 978-1-4799-3925-1.
[6] Liu Zhiguo, Kang Jincui, Li Yuan "A hybrid method of rough set and support vector machine in network intrusion detection" 2nd International Conference on Signal Processing System (ICSPS),IEEE 2010, ISBN:978-1-4244-6892-8.
[7] Zaman S, Karray F" Fuzzy ESVDF approach for intrusion detection system", International Conference on Advanced Information Networking and Applications(AINA), IEEE 2009, ISBN:978-1-4244-4000-9.
[8] V.N. Vapnik, "An Overview of Statistical Learning Theory", IEEE Transactions on Neural Networks, pp.10(5):988-999, 1999, ISSN :1045-9227.
[9] Xuehua Li, Lan Shu "Fuzzy Theory Based Support Vector Machine Classifier", Fifth International Conference on Fuzzy Systems and Knowledge Discovery, IEEE 2008, ISBN: 978-0-7695-3305-6.
[10] Tesfahun, Abebe, and D. Lalitha Bhaskari, "Intrusion Detection Using Random Forests Classifier with SMOTE and Feature Reduction." International Conference on Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), IEEE, 2013.
[11] Shah, Bhavin, and Bhushan H. Trivedi. "Optimizing Back Propagation Parameters For Anomaly Detection." IEEE-International Conference on Research and Development Prospectus on Engineering and Technology (ICRDPET). 2013.
[12] Shah, Bhavin, and Bhushan H. Trivedi. "Reducing Features of KDD CUP 1999 Dataset For Anomaly Detection Using Back Propagation Neural Network."Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on. IEEE, 2015.
[13] Shah, Bhavin, and Bhushan H. Trivedi. "Improving Performance of Mobile Agent Based Intrusion Detection System." Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on. IEEE, 2015.
[14] Pooja Champaneria, Prof. Bhavin Shah, Asst. Prof. Krunal Panchal, "Survey on Intrusion Detection System using Support Vector Machine", International Journal of Emerging Technology and Advanced Engineering (IJETAE), ISO 9001:2008 Certified Journal, Volume 4, Issue 12, December 2014, ISSN 2250-2459.
[15] Shah, Bhavin, and Bhushan H. Trivedi. "Artificial neural network based intrusion detection system: A survey." International Journal of Computer Applications 39.6 (2012).
[16] Cortes C. Vapnik V "Support-vector networks",1995 Machine Learning20, pp 273.
[17] A. Zadeh "Fuzzy Sets", Information and Control 8,pp.338-353, 1965.
[18] Mahesh Pal,"Multiclass Approaches for Support Vector Machine Based Land Cover Classification",tech report, National Institute of Technology, arXiv preprint arXiv:0802.2411 (2008).
[19] H. Yu and S. Kim, "SVM tutorial: Classification, Regression, and Ranking", Handbook of Natural Computing, pp.479-506, Springer, 2012.
[20] A. Zadeh "Fuzzy Sets", Information and Control 8,pp.338-353, 1965.
[21] Bhavin Shah, Bhushan Trivedi, "Dataset Normalization: For Anomaly Detection Using Back Propagation Neural Network".
[22] Ginny Mak, G. Ratzer, H. Vangheluwe, "The implementation of support vector machine using the sequential minimal optimization approach".
[23] "NEURAL NETWORKS" by Christos Stergiou and Dimitrios Siganos.
[24] http://scikit-learn.org/stable/modules/svm.html
[25] KDD Cup 1999 DATA, the UCI KDD Archive Information and Computer Science, University of California, Irvine, http://kdd.ics.uci.edu/databases/kddcup99/task.htmllast edited on November 30, 1999 at 10:46am