An Efficient Hybrid Classifier Model for Anomaly Intrusion Detection System

Asghar Ali Shah[1, 2], M. Khurram Ehsan[1], Kashif Ishaq[2], Zakir Ali[1], Muhammad Shoaib Farooq[2]

Bahria University Lahore Campus[1], University of Management and Technology[2]

Abstract

Ensuring security has always been a challenging problem for both customized network solutions and information systems. Intrusion Detection System (IDS) is playing a very important role to ensure security both in network solutions and information systems. Significant efforts has already been made and many efforts are underway to improve the IDS but still there are many short comings. This study proposed a model based on extensive survey to create an efficient hybrid classifier which is jointly based on feature selection, parameter optimization and classification. Feature selection is adapted to refine the area of interest by improving the accuracy of classification, then to optimize the parameters, genetic algorithm (GA) is the most appropriate technique to be used. Parameters optimization using GA also plays a remarkable role to improve classification using support vector machine (SVM). SVM is considered a suitable machine learning technique for classification of intrusions which are detected both in networks and information systems. Finally SVM will classify the observed activity as normal or attack using adopted linear or nonlinear techniques. The proposed solution paves a way to improve accuracy by efficiently detecting the intrusions within real time applications of network and systems.

Keywords:

IDS, Classification, SVM, GA, Attacks types

1. Introduction

To ensure high security of a network and information system is always a significant problem to be addressed continuously in both academia and industry. As network and system security concerns are increasing, many encryption schemes have been devised to address the said problem. Different firewalls both in software and hardware have implemented whereas a lot of access control methods are also introduced. But all of them are not well sufficient to ensure security. Assume that if password is weak, then system resources can be accessed easily by unauthorized users. Insider attack, malicious mobile code, insecure modems and error in configuration of firewalls are the threats which cannot be resolved completely by firewalls. Most of the attacks belong to insiders who misuse their privileges. Having above said problems, Intrusion Detection System (IDS) is an additional wall of security to resolve the issue. It has a key role in insuring security for both network and information system by detecting such malicious activity (intrusion) in a systematic manner [1]. A lot of work has been done to establish various IDSs including different techniques which are proposed to address the needs of existing infrastructure.

IDS can be classified into two different types based on where we look for intrusion. The details of which are discussed as following:

a. Host Based IDS: It provides the better visibility of all the occurrences on your critical system. It is used to monitor the activity on the system and also able to detect the threats coming from the network to the host.

b. Network Based IDS: Network based IDS monitors the voluminous activities coming from both outside and inside of the network.

While IDS can also be categorized based on detection as following:

a. Misuse Detection: Such IDS has a large database of previously known attacks. Where it compares each packet with its database of attacks, as a result an alert is generated by IDS if received packet matches any attack within database otherwise it proceeds [5]. On contrary it is completely blind to new attacks. So such IDS is efficient and accurate in order to detect

b. Anomaly Detection: If any abnormal activity is occurred within network or system and such activity is also not previously known, then the detection of such attack is known as anomaly detection. It is very complex and used to detect unusual activity. Such IDS is used mostly for detection of new attacks whereas false positive rate of alarms are very high

While malicious users of network and information systems can be categorized as external or internal, two different groups of intruders namely, external and internal intruders [3, 4].

a. External Intruders: When networks or information systems face the interruption from unauthorized users, such users are classified as external intruders. They use various penetration techniques to break into the system.

b. Internal Intruders: Authorized users within the networks or information systems with some restrictions are

Manuscript received November 5, 2018
Manuscript revised November 20, 2018

known as internal intruders. Usually they perform unauthorized activity within a network or information system.

It is really a difficult task to secure the network or information system ideally. Efficiency and access control has always inverse relationship. IDS monitors the activities in the network and information system and considered as defensive mechanism for security.

In this survey study intrusions detection using feature selection, parameter optimization and classification are discussed in details based on previous contributions done in perspective scenarios. Finally based on this extensive survey, hybrid approach is just proposed where feature selection, parameter optimization and classification are integrated jointly to improve the overall performance of IDS. IDS has good generalization capability and it is mostly used in real world applications [6], while an example of input data used for intrusion detection purpose is KDD CUP 1999 dataset [7].

2. Feature Selection

Appropriate feature selection helps to reduce training and as well as testing time of classifier to improve the performance and accuracy of classification. Better feature selection also reduces model complexity and computation cost. Feature selection is an activity which selects subset of near and relevant features [9] to create robust learning models actually by reducing the number of attributes. An analysis of machine learning techniques is performed to better specify their usage frequency [10]. In order to improve performance, it is required to devise such model with minimal number of features that discriminate normal from abnormal behavior accurately [11]. Different algorithms and machine learning techniques are therefore used to improve the feature selection process which is given as follows:

Genetic Algorithm (GA) is used for feature selection in [1, 12, 13, 14, 15, and 16]. The authors in [1] define three components as follows:

• Individual gene is initialized first
• Functional model is devised for evaluation purpose
•Specific function is then created for genetic operation

GA and SVM, both are combined for feature selection in [12]. Where Fitness function measures the fitness of chromosomes first. Lower fitness chromosomes are omitted then using GA. This process is repeated until the high fitness chromosomes are moved to the next generation and an individual good chromosome is to be detected.

In [13], GA and Balanced Iterative Reducing and Clustering (BRICH) using Hierarchies are used for feature selection. GA and BRICH help to reduce dataset using selection of relevant feature set. GA is used to select appropriate features from large dataset. BRICH clustering algorithm creates Clustering Features (CF) tree, which provides the brief representation of dataset.

Wrapper based feature selection is also used to extract relevant feature sets [14]. The proposed model has high detection accuracy 99.5% and fewer false alarms. Specifically PMU features are reduced from 41 to 6 attributes which results in 85% of space reduction for features. It also reduces training time and increases the performance of classification.

Feature selection is achieved using rough set theory which eliminates the redundant and useless features by applying GA [15].

This study consists of three phases such as dataset, design and evaluation which is described as follows:

- The NSL-KDD contains raw data in a tabular format where columns represent records and row represents attributes
- Redact is applied using GA based on the rough set model.
- The effect of feature selection using rough set model is then compared with other state of art feature selection algorithms

RSES simulator is used to calculate genetic redact based on the rough set model and Weka simulator is used to implement intrusion detection system.

In [16], the dataset is first labeled and GA is then used for feature selection on the new processed dataset. It provides set of metrics using GA which generates the more suitable intrusion detection results in a given environment. GA follows a few steps here as:

- Initialization: GA randomly selects a subset of chromosomes as initial population.
- Evaluation: Fitness function for each chromosome is estimated from initial population.
- Selection: The better fitness value of chromosome has a higher probability for selection. Roulette wheel gives a biased weight to each of the selected chromosomes on the fitness value.
- Cross Over: If pair of chromosome are mates with the probability Pc, then the last bit is interchanged with the cross over point. If pair of chromosome are not mate then it is unmodified.
- Mutation: Randomly modify the new population of chromosome to achieved diversity.

Particle Swarm Optimization (PSO) is chosen for feature selection in [17, 18, 19]. Intrusion Feature Selection Algorithm (IFSA) is proposed in [18], which is actually based on PSO. It actually explores the feature space and autonomously selects effective feature subset to reduce data attributes.

Fitness evaluation function is used in the selection of an attribute subset. The proposed algorithm is detailed as follows:

1. Load training dataset and set initial parameters
2. Randomly initialize velocity for each particle
3. Measure fitness value of each particle
4. Optimal location is to be selected
5. Update particle location and velocity

PSO chooses multiple set of attributes for optimal feature selection [19].

Principal Component Analysis is another technique which is used for feature selection as described in [20, 21, 22]. Kernel principal component analysis (KPCA) performs better than PCA due to better handling of nonlinearity in data [20]. PCA selects few features only by reducing the dimension with the help of linear transformation of dataset to feature space [21]. In [22], KPCA is adapted as a preprocessor for SVM to extract features. At first the features are mapped from higher dimension to a lower one, then GA is considered for further optimization of KPCA.

Feature selection is achieved through principle of Mutual Information (MI) in [23, 24]. Proposed schemes select features using filter ranking and Wrapper-based Improve Forward techniques.

Bees Algorithm (BA) is used for feature selection both in [25, 26]. In [25], feature selection is made using BA and Linear Genetic Programming (LGP), where LGP helps to improve better detection and low false alarm is achieved using BA. Membrane algorithm is proposed to enhance the Bee algorithm performance for better feature selections in [26].

SVM (FS-SVM) is proposed as feature selection algorithm for attributes reduction of sample data [27]. It gradually removes subordinate features from feature dataset.

In [29], Cuttlefish Algorithm (CFA) is used for features selection It removes both the redundant and irrelevant features which clearly will improve systems performance.

A toolbox (FEAST) in MATLAB used for intrusion detection based on feature selection. It provides results with better precision having least number of features in IDS applications [30].

K-mean clustering is used for feature selection too [31], where the features are selected based on intra class variance difference.

Three other standard feature selection techniques for the development of efficient and effective IDS are included Correlation based Gain Ratio, Feature Selection and Information Gain [60]. Naïve Bayes Classifier (NBC) is used perform the feature selection in preprocessing stage and then the output is given to Feature Vitality Based Reduction Method (FVBRM). Finally the output of three feature selection techniques and FVBRM is compared. The selected features are also discretized. Data Reduction, Data Integration and Data Cleaning techniques are also used in preprocessing.

Pulse Coupled Neural Network (PCNN) is used for feature selection. PCNN is a third generation artificial neural network used for feature selection [32].

Rough set theory in [33], used to extract relevant attributes and to classify the data packets. Rough set theory obtains the discernibility matrix for feature reduction. The discernibility matrix is simplified and help to obtain minimum number of attributes which is enough for classification of packets as normal or attack.

3. Parameter Optimization

Parameter optimization has a great impact on the performance of classification. Therefore a lot of work has been done for parameter optimization which clearly help to better detect intrusions in networks and information systems. In this survey some of the important techniques for parameter optimization are discussed.

GA is frequently used for the said purpose [20, 22, 34, 35, 36, 37, 38, 39, 40].

Optimization problem is addressed with the help of improved GA in [35]. It also introduced greedy algorithm to improve the efficiency of optimization algorithms. A mathematical model is devised and improved GA is also used to optimize the devised model. The proposed system answers the following three questions.

1. How to improve the reliability of the detection system resources constraints
2. How to minimize system resource under reliability constraints
3. How to construct a clear and complete mathematical model to address these two above concerns

In testing activities n-detection agents participated in given detection time. The reliability of each agent and its detection time is known.

Basic assumption:

1. There is no network congestion, transmission interruption and any other unforeseen circumstances.
2. Average detection time is known
3. Detection agent is at start and in working condition.

Multi-agent process optimization mathematical model.

A mathematical model of the above said problem is devised. It improves the efficiency and reliability of optimization keeping in mind the constraint of computational cost.

GA is also used for hyper-plane optimization [36]. In this approach, generally the parameters are partially optimized which can greatly affect the practical application of SVM.

In this study immune algorithm is introduced for SVM parameter optimization. SVM adapted the following procedure based on GA optimization.

1. System is initialized with SVM parameters
2. SVM functions behave as antigen for parameter optimization
3. Antigen and antibody are estimated using target function
4. An excellent antibody is recorded during the evolution process and the appearance of the cells is log down. Move to step 5 when the optimal parameters are found
5. Each antibody having its concentration and survival rate is estimated. Supply antibody selection and immune system
6. Update the new group. New group is created by selecting, recovering, mutating and then remove the rookie individual from the memory base, finally repeat from step 3

Chromosome structure can be expressed as

$Ch = (w1, w2,…., wf\_num, b, \xi\ 1,…., \xi s\_num)$

Chromosome parameters randomly generate initial group as

$$\left\{ \begin{array}{c} W_i \epsilon\ [-1, 1] \\ b\ \epsilon\ [-1, 1] \\ Ch= (w_1, w_2,…., w_{f\_num}, b, \xi_1,., \xi_{s\_num})\ \xi_j\ \epsilon\ [0,1] \\ i = 1, 2,… ,s\_num \\ j = 1, 2,… ,f\_num \end{array} \right\}$$

Genetic operators are selected using tournament selection, elitist selection, intermediate crossover and uniform mutation. The algorithm have to continue until it has encountered any of following condition:

- Fixed stagnation algebra
- Average adaptable and stable parameter function rate less than a fixed value

Improved GA with the following implementation steps are described in [39]:

- Initialize the parameters and binary coding
- Calculate the fitness of the individual
- Apply the elitist strategy
- Processing genetic manipulation
- Termination condition

The framework of intrusion detection model based on KPCA-SVM with GA [22] works as follows:

1. Capture data: Original Dataset
2. Normalized dataset
3. KPCA analysis: Feature extraction
4. Optimized parameters by using GA
5. Training data
6. Training of SVM with GA
7. SVM output: classify actions as normal or attack

PSO is more widely used algorithm after GA for parameters optimization which is discussed in [17, 27, 34, 40, 41, 42]. PSO-SVM is used for parameter optimization which helps for further efficient classification relevant attributes [34]. QPSO is proposed to tune properly and optimize the parameters of Radial Bases Function (RBF). The fitness value is evaluated and based on that obtained value the optimized parameter is achieved. RBF network is trained and then repeated until all parameters are tuned.
In PSO-SVM penalty parameters C and Kernel Function parameter g is optimized using improved PSO for SVM classification [40].
Flexible Neural Tree (FNT) selects best input variables and based on these input variables, better classifies intrusions [42]. It uses PSO for tuning parameters which improves the structure and then tunes parameters in efficient manner to obtain better optimization result. Different kernels are also used to optimize given parameters for classification [43]. RBF is used to optimize parameters in successive training and testing. A sigmoid kernel is used for classifier training and testing. Grid search is used for optimization of SVM parameters. Bees Algorithm (BA) is also used for optimization in [25].

4. Classification

Classification is widely used for intrusions detection. SVM is most appropriate technique used in classification as described in [1, 2, 34, 36, 28, 12, 44, 45, 37, 23, 20, 13, 14, 38, 46, 29, 39, 27, 43, 47, 31, 48, 17, 32, 49, 50]. In [1], authors proposed an enhanced SVM approach which is used for classification. Internet has mostly normal data patterns and anomalous traffic is relatively small. When the proposed SVM executed in soft margin, then it showed good performance and high detection rate than an unsupervised SVM. The reason is, proposed SVM has both supervised and unsupervised features and it is operated without the use of labels. The proposed SVM is described as following:

$$Yj (WT\O (xj)) \leq p - \xi \qquad 0 \leq p \equiv 1$$

A new ensemble classifier is proposed using bagging and results showed improvement in terms of accuracy [50]. The based classifier is built using RBF and SVM. Ten-fold cross validation is applied to the base classifier and classification accuracy is evaluated. To improve the

classification performance, bagging is also used with RBF and SVM.

In [51], hybrid algorithm is proposed which combine Modified Artificial Bee Colony (MABC) and Enhanced Swam Optimization (EPSO) for classification of intrusion, where KDDCup99 dataset is used for experiment. The performance is compared with other machine learning algorithms and found that it is improved in comparison to the state of the art work.

The framework of this study is given as follows:

- Data preprocessing: duplicate and redundant features are removed and the data is then processed for classification.
- Feature selection: Single Feature Subset Method (SFSM) and Radom Feature Subset Method (RFSM) are used to determine the feature subset.
- Hybrid classification: MABC and EPSO algorithms are combined for classification to obtain results with higher accuracy.

Hybrid classification approach is descried as following:

1. Initialization of parameters
2. Initialization of the colony
3. Evaluate the fitness value of each solution
4. If It is better than the previous one it change the old one with the new position otherwise it keep the previous solution
5. EPSO control the velocity
6. The best value is specified by MABC and EPSO
7. If termination condition is met then it terminates the process and produce the best solution otherwise reiterate to step 2

A new classifier Combining Support Vectors with Bee Colony (CSVBC) is proposed [48] which combine SVM and Bee Colony. It is proposed to achieve high performance of IDS. A modified binary classifier CSVBC is devised to differentiate between normal and abnormal data.

In [53], Multi Class Batch Support Vector Machine (MCBSVM) is proposed which differentiate hyper plane quickly into multiple classes and find the right class position.

An online one-class classification approach is proposed [54] which is based on the mahalanobis distance. One-class problem is defined by two concentric hyper spheres enclosing the support vectors of the description. The classifier is updated after each iteration, which is tested on real dataset by providing different types of attacks. The quadratic mahalanobis distance is the feature space estimated between each sample. The one class problem allows to differentiate outliers from the support vectors.

The proposed system architecture [58] is consisted of three phases detailed as following:

Preprocessing Phase: Differentiate dataset: training set and testing set. Each record has only six features of the KDD dataset.

Training Phase: (QVICA with EDA) is trained to learn for classification of data into either normal or attack. The algorithm initializes quantum anti body population and quantum vaccine population. The following processes are used in this stage.

- Initialization
- Vaccine decoding
- Vaccine selection and vaccination
- Vaccine sampling
- Clonal selection

Testing Phase: The proposed system is then tested for evaluation using training data samples.

In [59], it is stated that SVM is the mostly used and considered to be the best machine learning technique for classification of intrusion detection

The summary of survey are detailed in table 1 & 2.

Table 1: Articles written for a specific category

Categories	Articles
Feature Selection	[1, 2, 12, 13, 14, 15, 16, 17, 18 ,19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 32, 33, 35, 43, 47, 60]
Parameters Optimization	[20, 22, 25, 31, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43]
Classification	[1, 2, 12, 13, 14, 17, 19, 20, 21, 22, 25, 27, 28, 29, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 61]

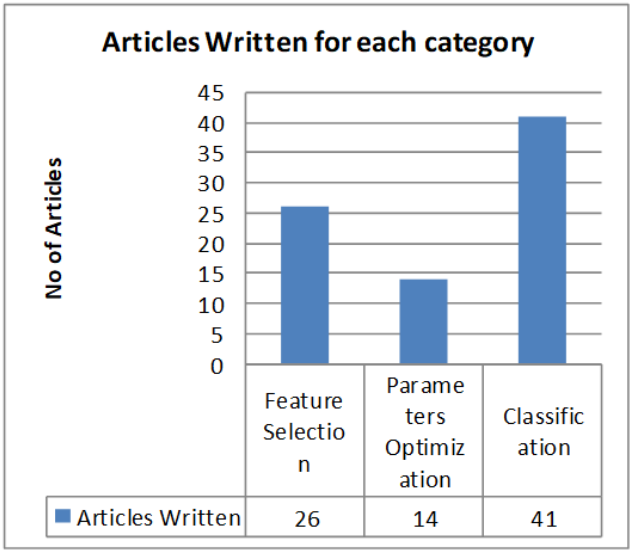

Fig. 1 Articles written for a specific category

Table 2 shows a specific machine learning technique for a given category and it is also highlighted in figure 2.

Table 2: Articles written for machine learning techniques

Machine learning Techniques/ categories	Feature Selection	Parameters Optimization	Classification
SVM	[27]		[1, 2, 12, 13, 14, 17, 20, 23, 25, 27, 28, 29, 31, 32, 34, 36, 37, 38, 39, 43, 44, 46, 47, 48, 49, 50]
GA	[1, 2, 12, 13, 14, 16, 40]	[20, 22, 34, 35, 36, 37, 38, 39, 40]	
PSO	[17, 18, 19]	[17, 27, 34, 40, 41, 42]	[51]
IGR	[28]		
MI	[23,24]		
PCA	[20, 22, 43]		
C4.5			[43]
C5			[29]
CFA	[29]		
MABC			[51]
Supervised and Outlier Method			[21]
FEAST Toolbox	[30]		
CART			[52]
Grid Search		[43]	
LGP	[25]	[25]	
BA	[25, 26]	[25]	
K-mean Clustering	[31]		
NBC	[60]		
BC			[48]
MCBSVM			[53]
MC	[26]		
FROAFTN			[19]
Mahalanobis Distance			[54]
RBF			[41, 50]
PCNN	[32]		
FNT			[42]
Rough Set Theory	[33]		
Rosetta Tool			[33]
BV-TCAM			[55]
GPBDT			[56]
SOM			[57]
QVICA			[58]

If we summarize the survey briefly, GA is widely used for feature selection and parameter optimization while in most of classification problems SVM is mostly used for real time IDS application scenarios.

5. Proposed Model

As discussed in detail, many researchers have adapted different approaches to devise efficient IDS. They are broadly based on feature selection, parameter optimization and classification techniques or used any hybrid techniques to resolve the security issues within network or information systems. But still the performance and accuracy of intrusions classification is under question. Based on this survey, a model is proposed for developing an efficient anomaly based IDS. The proposed model is consisted of following three phases:

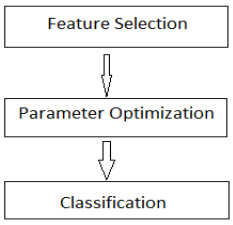

Fig. 2 Proposed Model

The purpose of such threefold proposed model is to obtain the better classification of intrusions having prior information of relevant features using feature selection and optimized parameters using parameters optimization techniques respectively. As the proposed model focuses on three aspects such as feature selection, parameters optimization and classification as a hybrid approach to address the issue of efficient classification of intrusions within network and information systems.

A hybrid classifier is proposed that integrates SVM and GA. Each learning classifier has different role. Multiple classifiers are combined to increase performance. GA is adapted to optimize all the selected parameters of SVM and then SVM is used for efficient intrusion detection. Parameters optimization plays a key role for SVM classification. It affects the distribution of samples in the feature space which can lead to over learning or less learning. So true optimization can also improve SVM performance and classification accuracy under the assumed constraints. Therefore, we use GA for optimizing the parameters of SVM. SVM is then adapted for intrusion detection because it is classifier based on margin. SVM algorithm then classify it as an attack or normal activity using linear or no linear chosen technique. The proposed IDS model is used for anomaly with better accuracy and less convergence time. The architecture of the proposed solution is given in figure 2.

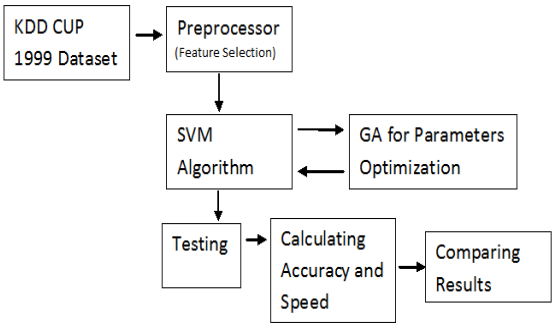

Fig. 3 Architecture of proposed solution

NSL KDD CUP 19999 Dataset: It is the initial dataset to be observed

Pre-processor: Reduce data by dimensionality reduction and removal of outlier using feature selection techniques.

GA for parameters optimization: GA is then used for parameters optimization. Parameters optimization has a great importance in performance of SVM

SVM algorithm: Finally the input dataset is classified to multiple categories for efficient intrusion detections.

Calculating Accuracy and Speed: It will measure the accuracy and computational speed of classification. The accuracy metric given in table 3 based on simple similarity measure.

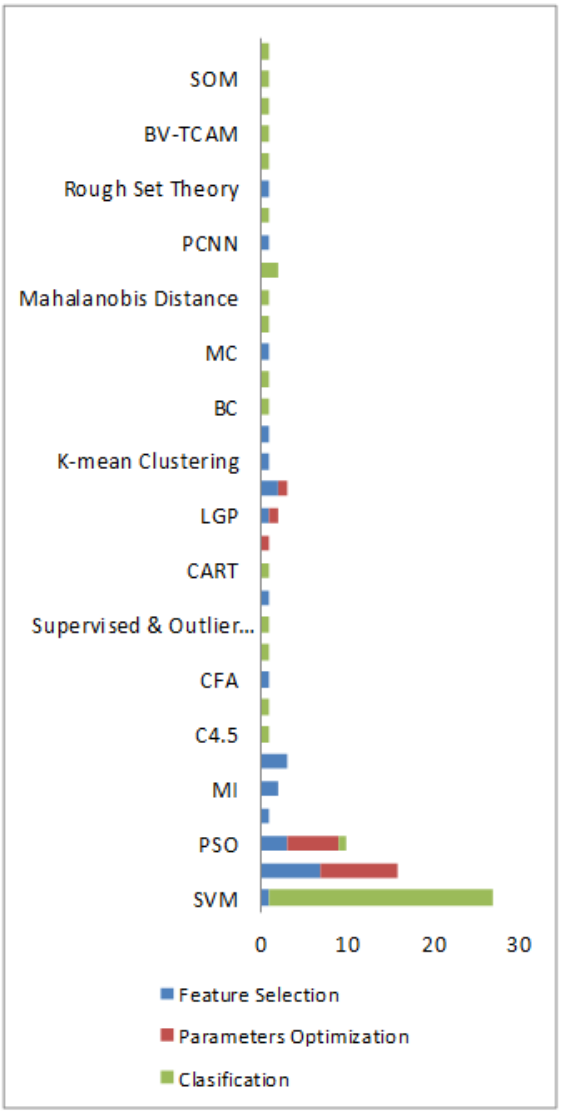

Fig. 4 Articles written for machine learning techniques with different categories.

Table 3: Accuracy Matrix

T/F	Positive	Negative
Positive	a	b
Negative	c	d

Accuracy= (a + b) / (a + b + c + d)

Classification of attacks: SVM is then classified the observed data that either it is normal activity or an attack. If it is an attack then it will be divided into further four categories which can be R2l, Probe, U2R and DOS. Classification of attacks is given in figure 2.

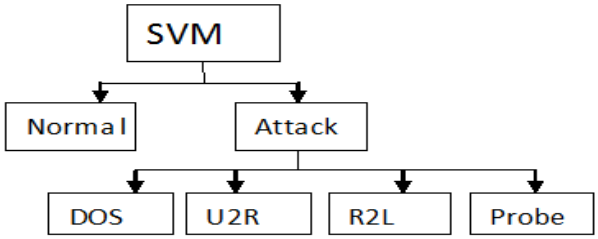

Fig. 5 Classification of four types of attacks

6. Conclusions

A detail survey study is conducted regarding intrusion detection system (IDS), it is concluded that in studied state of art work, IDS are implemented widely using either one of them including feature selection, parameter optimization and classification or any hybrid approach is adopted. A hybrid classifier model is then proposed based on concluded survey study. The proposed threefold model helps to improve the classification of intrusions having useful prior knowledge from feature selection and parameter optimization. The feature selection approach is used to reduce the attributes of dataset by narrow down the relevant features in hand. GA is used for parameter optimization. Both having relevant feature dataset and usage of parameter optimization technique then increase the performance of classification algorithm. SVM is used finally for classification of observed dataset. For experimental verification, NSL KDD CUP 1999 dataset is to be recommended.

References

[1] A.K. Jones, and R. S. Sielken. “Computer System Intrusion Detection: A Survey”. Technical Report. Department of Computer Science, University of Virginia, Charlottesville, Virginia, 2000.

[2] R. G. Bace, : “Intrusion Detection”. Macmillam Technical Publishing (2000).

[3] R. Gong, H., M. Zulkarnine, and P. Abolmaesumi, (2005); A Software Implementation of GA based approach to network intrusion detection, http: //www.cse.msu.edu/~cse848/2011/student papers/Tavon Pourboghrat.pdf.

[4] S. S. Kandeeban and R. S. Rajesh, (2007); GA for framing rules for intrusion detection, J. Comp. Sci and Security., 7(11). ISSN: 1738-7906, PP. 285-290

[5] V. N. Vapnik, the nature of statistical learning theory. Springer – Verlag, New York. NY, 1995.

[6] A. Majid, A. Khan, A. M. Mirza, Combining support vector machines using geneticprogramming, Int. J. Hybrid Intell. Syst. 3 (2) (2006) 109–125.

[7] S. J. Stolfo, W. Fan, A. Prodromidis, et al., KDD Cup 1999 Data [EB/OL] (2011)http://kdd.ics.uci.edu/ databases/ kddcup99/kddcup99.html.

[8] PredictionWorks: Data Mining Glossary. (n.d.). PredictionWorks. Retrieved February 11, 2011, from http://www.predictionworks.com/glossary/index.html

[9] O. S. Soliman and A. Rassem, “A network intrusion detection system based on a quantum bio inspired algorithm,” International journal of Engineering Trends and Technology, vol. 10, no. 8, pp. 371-379, Apr. 2014.

[10] A. A. Shah, M. S. H. Khayal and M. D. Awan, “Analysis of machine learning techniques for intrusion detection system: a review,” International Journal of Computer Applications, vol. 119, no. 3, pp. 19-29, June 2015.

[11] G. H. Johan, R. Kohavi and K. Pfleger, “Irrelevant feature and the subset selection problem,” Proc. Of the 11th int. Conf. on Machine Learning, Morgan Kaufmann Publishers, pp. 121-129, 1994.

[12] X. Q. Wang, “Study on Genetic Algorithm Optimization for Support Vector Machine in NetworkIntrusion Detection,”AISS: Advances in Information Sciences and Service Sciences, vol. 4, no. 2,(2012), pp. 282-288.

[13] M. A. Ambusaidi, X. He, Z. Tan, P. Nanda, F. Lu and U. T. Nagar, “A novel feature selection approach for intrusion detection data classification,” Center for Innovation in IT Services &Applic., Sydney, Australia 2014.

[14] F. Kuang, W. Xu and S. Zhang, “A novel Hybrid KPCA and SVM with GA model for intrusion detection ,“Journal of Applied Soft Computing, vol. 18, pp. 178-184, 2014.

[15] A. N. Devi and K. P. M. Kumar, “Intrusion detection System based on genetic-SVM for DoS attacks,” International Journal of Engineering Research and General Science, vol. 3, no. 2, pp. 107-113, 2015.

[16] G. Wang, S. Chen and J. Liu, “Anomaly based intrusion detection using multiclass-SVM with parameters optimized by PSO,” International Journal of security and its Applications, vol. 9, no. 6, pp. 227-242, June 2015.

[17] M. Gupta and S. K. Shrivastava, “Intrusion detection system based on SVM and Bee Colony,” International Journal of Computer Applications, vol. 111, no. 10, pp. 27-32, Feb. 2015.

[18] V. Manekar and K. Waghmare, “Intrusion detection system using support vector machine (SVM) and particle swarm optimization (PSO),” International Journal of Advanced Computer Research, vol. 4, no. 3, pp. 808-812, Sep. 2014.

[19] K. I. Rufai, R. C. Muniyandi and Z. Othman, “Improving bee algorithm based feature selection in intrusion detection system using membrane computing,” Journal of Networks, vol. 9, no. 3, pp. 523-529, Mar. 2014.

[20] K. Atefi, S. Yahya, A. Y. Dak and A. Atefi, “A hybrid intrusion detection system based on different machine learning algorithms,”Proceeding of the 4th ICOCI, Sarawak, Malaysia, 2013, pp. 312-320.

[21] P. Amudha, S. Karthik and S. Sivakumari, “A hybrid swam intelligence algorithm for intrusion detection using significant features,” The Scientific World Journal, vol. 2015, no. 5, May 2015.

[22] P. Nader, P. Honeine and P. Beauseroy, “Online one-class classification for intrusion detection based on the mahalanobis distance,” In Proc. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Belgium, pp. 567-571, Apr. 2015.

[23] U. S. Aswal, J. Bhatt, S. K. Chauhan, " Support vector machine based intrusion detection system," International Journal of Computer science and technology (IJCST), vol. 4, issue 2, pp. 84-87, 2013.
[24] N. Kausar, B. B. Samir, I. Ahmad and M. Hussain, "Efficient intrusion detection system based on support vector machines using optimized kernel function."Journal of Theoretical and Applied Information Technology, vol. 60, no. 1, pp. 55-63, Feb. 2014.
[25] J. Song, Z. Zhu and C. Price, "Feature grouping for intrusion detection based on mutual information," Journal of Communication, val. 9, no. 12, pp. 987-993, Dec. 2014.
[26] V. Jaiganesh, D. P. Sumathi and A. Vinitha, "Intrusion detection system using multiclass batch algorithm," International Journal of Advanced Research in Computer and Communication Engineering, vol. 4, no. 2, pp. 60-62, Feb. 2015.
[27] C. Zhang and Z. Fang, "A new distributed intrusion detection system model based on SVM in wireless Mesh networks," Journal of Information & Computational Science, vol. 12, no. 2, pp. 751-759, june 2015.
[28] S. Liu, L. Yu and Y. Fang, "Application of improved genetic algorithm in reliability optimization of multi-agent intrusion detection," in 2nd International Conference on Electronic & Mechanical Engineering and Information Technology (EMEIT)., 2012, pp. 2317-2320.
[29] M. Solanki and V. Dhamdhere, "A hybrid approach for intrusion detection using data mining," International Journal of Innovative Research in Science, Engineering and Technology, vol. 4, no. 7, pp. 5588-5595, Jul. 2015.
[30] J. A. Jeyanna, E. Indumathiand Dr. D. S. Punithavathani, "A network intrusion detection system using clustering and outlier detection," vol. 3, no. 2, pp. 975-982, Feb. 2015.
[31] V. K. S., D. N. Kasthuri and V. R. Saraswathy, "Enhancement of feature selection technique for network dataset," International Journal of Advanced Research Trends in Engineering and Technology (IJARTET), vol. 2, no. 8, pp. 22-26, Feb. 2015.
[32] H. Sheth, Prof. B. Shah and S. Yagnik, "Oprimizing and tuning RBF parameters using QPSO algorithm for anomaly detection in network intrusion detection system," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, no. 6, pp. 641-645, June 2015.
[33] M. Govindarajan, "Hybrid intrusion detection using ensemble of classification methods," I. J. Computer Network and Inforamtion Security, vol. 2, no. 7, pp. 45-53, 2014.
[34] T. Shon, Y. Kim, C. Lee and J. Moon, "A machine learning framework for network anomaly detection using SVM and GA," In Proc. IEEEWorkshop on Information Assurance and Security, New Yark, 2005, pp. 176-183.
[35] T. Shon and J. Moon, 2007, "A hybrid machine learning approach to network anomaly detection," Inf. Sci.,vol. 177, no. 18, pp. 3799-3821, 2007.
[36] Y. Ren and G. Bai, "Determination of optimal SVM parameters by using GA/PSO,"Journal of Computers, vol. 5, no. 8, pp. 1160-1168, 2010.
[37] F. Amiri, et al… "Mutual information-based feature selection for intrusion detection systems," Journal of Network and Computer Applications, vol. 34, pp. 1184-1199, 2011.
[38] M. Bisen and A. Dubey, " An intrusion detection system based on support vector machine using hierarchical clustering and genetic algorithm," The Standard International journal, vol. 3, no. 1, pp. 21-25, 2015.
[39] M. S. Rani and S. B. Xavier, "A hybrid intrusion detection system based on C5.0 decision tree and one class SVM with CFA," International Journal of Innovative Research in Computer and Communication Engineering, vol. 3, no. 6, June 2015.
[40] A. Shrivastava, M. Baghel and H. Gupta, "A novel hybrid feature selection and intrusion detection based on PCNN and support vector machine," International Journal of Computer Technology and Applications, vol. 4, no. 6, pp. 922-927, 2013.
[41] V. Singh and H. Arora, "Network intrusion detection using feature selection and PROAFTN classification," International Journal of Scientific and Engineering Research, vol. 6, no. 4, pp. 466-472, Apr. 2015.
[42] Z. Jian-hua and L. Wei-hua, "Intrusoin detection research based on SVM and Intelligence algorithm," International Journal of Advancements in Computing Technology, vol. 4, No.16, pp. 445-452, Sept. 2012.
[43] F. J. Aparicio-Novarro, K. G. Kyriakopoulos and D. J. Parish, "Automatic dataset labeling and feature selection for intrusion detection system," In Proc. IEEE Military Communication Conference MILCOM pp. 46-51, 2014.
[44] A. Dastanpour, R.A.R. Mahmood, "Feature selectionbased on genetic algorithm and support vector machinefor intrusion detection system,"In: The Second International Conference on Informatics Engineering & Information Science, Nov.1214, 2013, pp. 169-181.
[45] J. Jha and L. Ragha, "Intrusion Detection System using Support Vector Machine,"International Journal of Applied Information Systems, New York, USA, 2013.
[46] S. Bordbar and M. T. B. Abdulah, "A feature selection based on rough set for improving intrusion detection system," The International Journal of Engineering and Science (IJES), vol. 4, no. 7, pp. 54-60, Jul. 2015.
[47] J. Patel and M. K. Panchal, "Effective intrusion detection system using data mining technique," Journal of Emerging Technologies and Innovative Research (JETIR), vol. 2, no. 6, pp. 1869-1877, June 2015.
[48] H. M. Tahir, W. Hasan, A. M. Said et al, "Hybrid Machine learning technique for intrusion detection system," In Proc. International Conference on Computing and Informatics, Instambul, Turkey, pp. 464-472, Aug. 2015.
[49] F. Kuang, W. Xu , S. Zhang, Y. Wang and K. Liu, "A novel approach of KPCA and SVM for intrusion detection," Journal of Computational Information Systems, vol. 8, no. 8, pp. 3237-3244, 2012.
[50] L. Khan, M. Awan and B. Thuraisingham, "A new intrusion detection system using support vector machines and hierarchical clustering," The VLDB Jouranl, vol. 16, pp. 507-521, 2007.
[51] M. S. Rani and S. B. Xavier, "A hybrid intrusion detection system based on C5.0 decision tree and one class SVM," International Journal of Current Engineering and Technology, vol. 5, no. 3, June 2015.

[52] M. P. Fabio, D. L. H. C. Eduardo and D. L. H. M. Alexis, "Application of FEAST (Feature Selection Toolbox) in IDS (Intrusion Detection System)," Journal of Theoretical and Applied Information Technology, val. 70, no. 3, pp. 579-585, Dec. 2014.
[53] N. R. Patil, A. R. Wadhawane, N. V. Sarode and S. S. Rathod, "Identification of reduced features for intrusion detection system using naïve bayes classifier," International Journal of Engineering and Technical Research, vol. 3, no. 4, pp. 36-40, Apr. 2015.
[54] L. Tong and Q. Wu, "Intrusion feature selection algorithm based on particle swam optimization," International Journal of Computer Science and Network Security, vol. 14, no. 12, pp. 40-44, Dec. 2014.
[55] Y. Chen, et al…, "Hybrid flexible neural tree based intrusion detection systems," International Journal of Intelligent Systems, vol. 22, pp. 337-352, 2007.
[56] N. Gupta, N. Singh, V. Sharma, T. Sharma and A. S. Bhandari, "Feature selection and classification of intrusion detection system using rough set," International Journal of Communication Network Security, vol. 2, no. 2, pp. 20-23, 2013.
[57] H. Song and J. W. Lockwood, "Efficient packet classification for network intrusion detection using FPGA," California, USA, Feb. 2005.
[58] S. Rathore, Prof. A. Saxena and Dr. M. Manoria, "An efficient grid partition based classification algorithm for intrusion in KDDCup 99," International Journal of Computer Science and Information Technologies, vol. 6, no. 3, pp.3073-3076, 2015.
[59] A. A. A. Ahrabi et al.., "A new system for clustering and classification of intrusion detection system alerts using self-organizing maps," International Journal of Computer Science and security, vol. 4, no. 6, pp. 589-597, 2009.
[60] S. R. Hasani, Z. A. Othman and S. M. M. Kahaki, "Hybrid feature selection algorithm for intrusion detection system," Journal of Computer Science, vol. 10, no. 6, pp. 1015-1025, 2014.
[61] N. Parati and S. Potteti, "Intelligent Intrusion detection system using SVM and Genetic Algorithm (SVM-GA)," International Journal of Science and Applied Information Technology (IJSAIT), vol. 4, no. 2, pp. 01-05, 2015.