Classification, Detection and Prevention of Network Attacks Using Rule Based Approach

Wrushal K Kirnapure[1], Arvind R. Bhagat Patil[2]

[1]Student, [2]Professor, *Dept. of Computer Science Engineering, YCCE college, Nagpur, India*

---***---

Abstract - *Intrusion detection systems classify normal and abnormal activities. Categorization can be used to ease the task of intrusion detection systems. Many machine learning algorithms have been developed to automate the task of classification and categorization. Hence we propose a machine-learning based classification algorithm against network attacks, while minimizing misclassification. The goal of Intrusion detection framework (IDS) is to apply all the available information in order to distinguish the attacks as either by outside programmers or abuse of insiders. Network intrusion detection design and development follows diverse arrangement models. The most frequently used machine learning algorithms in these models are Support Vector Machine (SVM) and Ant Colony. Considering there weaknesses and strengths of both a blend of the two algorithms is developed for intrusion detection system in this paper. The combined approach outperforms both SVM and ant colony when used alone. A standard benchmark data-set which contains variety of intrusions simulated in military network environment, the KDD Cup 99 is used. This dataset is trained using Genetic algorithm. The rules for the SVM classifier are generated once the dataset is trained. The developed approach will be evaluated using parameters: detection rate, false alarm rate.*

1. INTRODUCTION

With the ever increasing traffic on networks, increasing complexity of attacks and drastic increase in number of networks everyday none of the present day stand-alone intrusion detection systems are capable of meeting the high demands for a system that has a very high detection rate and an extremely low false alarm rate. Also, most of the IDSs available in literature show different preference for detecting a certain class of attack with improved accuracy while performing moderately for the other classes of attacks. Hence there is a need to develop a system that detects a wide array of attacks and has a low false alarm rate and high detection rate. Support Vector Machines have been widely accepted as a powerful data classification method. On the other hand, the Self-Organized Ant Colony Network has been shown to be efficient in data clustering. Hence we have combined both in order to take advantage of their strengths while avoiding their weaknesses. A major challenge in developing IDSs is to realize real-time detection in high-speed networks. There are two important issues for this problem. First in order to reduce the cost of deploying a model, we must be able to minimize the cost of clean data that is used by the data mining process that is to reduce the amount of data required to train a classifier and reduce the training time. The machine-learning-based SVM method is a suitable for such an approach and can be trained with little volume of data. Second, when new information is added into a system, updated of the old model is required immediately, to ensure that the system is properly protected against modern attacks.

Thus, a mechanism is needed to generate an adaptive model, that can be updated, by integration of the new information that is gathered about the modern attacks into the old model of intrusion detection.

2. RELATED WORK

In paper [1] the principle of multi scale is introduced. The proposed work involves use of graphs of different scales which are extracted by isolating the video diagrams into squares of different sizes to achieve more human eye flexibility. Then the spatial components, specifically luminance, chrominance and surface, are isolated particularly by using discrete cosine change coefficients while the transient information is expelled from the development vectors to outline the heuristic cross sections. Next, the heuristic lattices are used as an element of the insect province improvement handle. Each heuristic cross section is used to coordinate the ants in the estimation and the ants store pheromone on the graph. The pheromone is overhauled through choking and vanishing, in this way surrounding spatial/transient saliency maps are generated. Finally, the spatial and transitory saliency maps of each scale are merged through adaptive mix, and maps of different scales are interwoven through direct mix. Since the model is created using information in short or tightly packed range, the decompression technique is avoided to save extra time and made suitable for capturing recorded transmissions on the system. Besides, the proposed method has been applied on a couple video databases with progressions in various scenes. Through examinations it has been observed that in both quantitative evaluation scores and normal visual effects, the results in this paper show a improved performance in contrast to other separation procedures surveyed in this paper. In paper [2] the authors present a novel approach called ACCMLF, which joins subterranean insect province batches with multilevel framework to reduce the runtime in the unlimited scale PPI systems. Here they have used a planning system to coarsen the principal enormous scale PPI effects, and get a compact PPI system. Then the insect state packing count is utilized to assemble and arrange. After this, the bundling outcome of exceptional framework is extracted through de-coarsening and then refinement is done to keep a separation so as to avoid falling into the neighboring area. Analysis of some large scale systems exhibits that the detection rate of ACC-MLF has significant improvements as opposed to ACC-FMD, and ACC-MLF can enhance gathering this is realized by performing some evaluation and estimations and differentiating the ACC-FMD, MCODE, MINE and Core computations. This framework for intrusion detection systems using sensitive figuring can be used to extend general framework security into specific framework security. In the paper [4] authors present the Mobile Network Defense (MND).It is a lightweight intrusion detection framework. MND is naturally inclined to be processed over masses of ants, giving it many advantages over other intrusion detection systems. In MND each subterranean insect in the virtual state can remember one specific metric of the current state of a computer. In mix, the delayed consequences of these fundamental tests can demonstrate specific assaults, while the dynamic method for the MND offers execution benefits over the standard static setup. This paper will show how the naturally organized MND offers a 34% change in detection time over other administrator based frameworks, and gives more successful intrusion detection arrangement than a static model with respect to CPU utilization, making the framework appealing for utilization in transversely over many sorts of mobile phones. The system in [5] fetches and shapes pictures at different diminishing levels through one or other camera units watching certain area(s) by method similar to a local area network(LAN) and is capable of combining information from different camera units to get a accurate and precise decision. It can be set up to capture certain kind of intrusions, for example individuals walking by, a social occasion for a group of walkers, vehicles, pets, etc .It can also reduce the number of false alerts due to other non-alarming intrusions. As a relevant examination, recognition of individuals either walking or on a vehicle in an observation domain is done using the proposed approach. This vision-based intrusion detection approach contains two essential steps: establishment subtraction based hypothesis generation (HG) and appearance based hypothesis evaluation (HV). HG estimates possible threats (intrusions), and HV checks those hypotheses using a Gabor channel for highlight extraction and bolster vector machines (SVMs) for portrayal. The framework has been analyzed in an unconstrained outdoor condition, providing good overall performance. The work in [3] presents an advanced approach in a direction where the function of the IDS is to handle or address a specific piece of the network attacks ordinarily recognized at port 7 in UDP. Port ranges in UDP speak to a sizable piece of the Internet development and almost little research depicts security in UDP port scan activity. To meet the new challenges posed by modern attacks other security issues in the continually expanding web field, this paper shows a computationally smart intrusion detection system using swarm intelligence techniques, particularly insect province change and ant colony optimization, to mitigate the effect of attacks in UDP. The principal purpose of the study in[6] is presenting port yields. This work focuses on creating a DOS attack intrusion detection framework by using support vector regression and improving this estimation by merging two computations of subterranean insect state and firefly. The firefly estimation changes ants' positions by doing a local hunt and firefly execution would be redesigned by reducing subjective parameters. Standard KDD Cup 99 course of action has been used for the evaluation of intrusion detection framework. This course of action consolidates 41 properties among which 9 properties have been picked. The proposed framework has 99.57% accuracy. The work in [7] is motivated due to the growing need for an IDS which should be capable of detecting intrusions of a wide variety at a faster rate and also provide better security with the ultimate objective of securing the networks. Existing models of intrusion detection frameworks (IDS) have made significant advances both in detection rate and security however their abilities suffer due to multilevel classes of attacks and intrusions with another problem which is the large amount of time taken for training the classifiers as well as getting the classifiers ready for execution . These drawbacks are avoided by developing hybrid models that combine the distinctive characteristics of single classifiers that is their strengths meanwhile avoiding their inadequacies for better execution. In this paper, a relationship of such composite models is presented. The objective is to choose their strengths and separate their weaknesses. In this way, an investigative gap is developed for more gainful intrusion detection models. In[8] this paper, the authors have presented a conveying gathering framework which does not require any prior model information. Specifically, at each of the sensor in the network, diverse two fold support vector machine (SVM) based classifiers are used and each classifier is set up to identify one class from the rest. At the center node for the sensor's, the Dempster-Shafer theory is realized to satisfactorily solidify the affirmation from all SVMs with appropriately computed fundamental probability assignments. The final conclusion is made by choosing the class with the most elevated conviction. Hypothetical performance expectation strategies are proposed for the planned arrangement framework. Through tests on a fabricated dataset and the benchmark 1999 KDD intrusion detection dataset, it is shown that the feasibility of the evaluation system and the commonness of the proposed framework is better in comparison to the customary Bayesian cost based combination in this particular situation. In the work presented in paper[9], memory standard and insect state ant colony computation are merged and integrated in an intrusion detection framework. The strategy for controlling pheromone used as a piece of subterranean insect settlement count is associated to imitate the memory method of human cerebrum, and the possibility of pheromone is progressed. The systems of recalling and neglecting are sensibly interpreted through the development and decreasing of pheromone, and in the count, new memory computation is molded to finish the methodology of holding and disregarding in the wake of considering the impact of bizarre qualities and their weights can have on the pheromone. Cases exhibit that the computations performed are fit for recognizing recalling and neglecting systems and growing the scope and self adaptability of IDS. This paper[10] proposes a novel intrusion detection approach by applying subterranean insect state change for important and highlighting decisions and SVM for detection. The intrusion segments are addressed as diagram ere center points, with the edges between them meaning the inclusion of the accompanying component. Ants cross through the graph to incorporate centers until the end standard is satisfied. The fisher detachment rate is held as the heuristic information for ants' traversal. In order to abstain from training endless classifier, the base or minimum square based SVM estimation is employed. At the start, the SVM is set up in light of grid chase system to get partition work using the planning data in perspective of all segments available, represented by network data. In paper [11] , the authors present a comparative survey of different methodologies for IDS. The issues in the genetic algorithm are discussed. An Intrusion Detection framework (IDS) based Hybrid Evolutionary Neural Network (HENN) using the genetic algorithm has been discussed. The authors have also concluded that the proposed framework improves detection rate and correctness rate. In paper [12], the authors have proposed an Intrusion detection framework for mobile ad-hoc networks (MANETs), as most of the currently available IDS are applicable to fixed networks. The proposed approach combines the genetic algorithm and artificial immune system. The proposed approach is called GAAIS Enactment limit, and life expectancy depend on hereditary calculation (GA) and manufactured investigation. The capacities of dynamic IDS to resist and detect several types of attacks such as Flooding, Black hole, Neighbor, Rushing and Wormhole is also presented and it has been demonstrated with reasonable experimentations that the proposed approach is efficient in comparison to other similar approaches. For incremental learning on new information, and element intrusion detection in AODV-based MANETs by adjusting to novel information is also demonstrated. GAAIS can adjust to network topology changes utilizing two methods : partial and total.

This paper[13] presents and researches the structure and build up of the chromosomes utilizing the conduct of dynamiCS that is dynamic clonal selection algorithm. The work is focused toward the continuously changing systems, adding dynamic learning abilities to it using the dynamiCS an element clonal choice, and allowing for the correct categorization of self patterns and also predicting new patterns of non-self. Transformations are done for determining calculations, intended to have such fuzzy govern sorts that organize assault information based on properties of self-adjustment. The impacts of three important framework parameters: tolerisation period, activation threshold and life span are explored. In[14] this paper the authors have proposed a transformative algorithm to enlist fuzzy categorization rules. The computation of rules for categorization utilizes an ant colony optimization based on nearby searcher to enhance the nature and quality of the eventual fuzzy categorization framework. The proposed calculation is performed on intrusion detection as a high-dimensional categorization problem. The experiments done by the authors demonstrate that the realized transformative ACO-Based algorithm is equipped for delivering a dependable fuzzy rule based classifier for intrusion detection.

To classify network activities (in the network log) as normal or abnormal while minimizing misclassification .To defend computer systems from various cyber attacks and computer viruses. To balance the performance of IDS in terms of efficiency and accuracy.

Intrusion Detection Systems (IDSs) are designed to defend computer systems from different cyber attacks and computer viruses.

I. IDSs build effective classification models or patterns to difference between normal behaviors from abnormal behaviors that are represented by network data.

II. To classify network activities (in the network log) as normal or abnormal while minimizing misclassification .To defend computer systems from various cyber attacks and computer viruses.

III. To balance the performance of IDS in terms of efficiency and accuracy.

3. PROPOSED SYSTEM

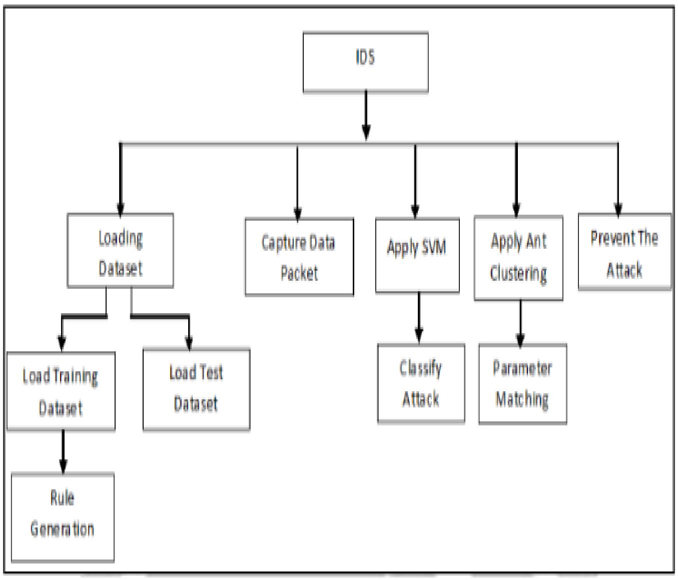
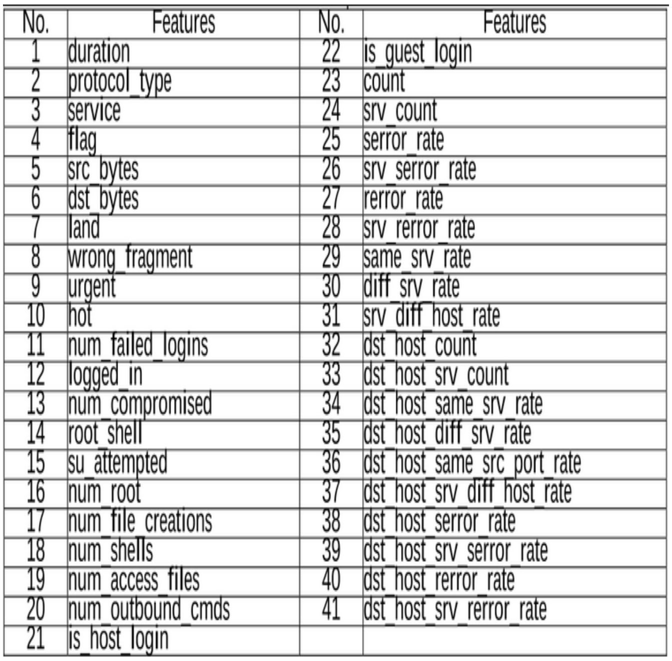


Fig.1. Architecture of Proposed work

Load Dataset

In this phase training dataset get loaded. There are many dataset but we are going to use KDDCup99 and NSLKDD. It will be given as input to the system and then rules are get generated and this rules will done the packet formation parameter.KDD99 is standard dataset for evaluation of data mining based IDSs. In KDD99, the data records of attacks fall into four main segments: Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R), and Probing. In KDD99, each record is described with the help of 41 features.

Table 41 parameters of packet



No.	Features	No.	Features
1	duration	22	is_guest_login
2	protocol_type	23	count
3	service	24	srv_count
4	flag	25	serror_rate
5	src_bytes	26	srv_serror_rate
6	dst_bytes	27	rerror_rate
7	land	28	srv_rerror_rate
8	wrong_fragment	29	same_srv_rate
9	urgent	30	diff_srv_rate
10	hot	31	srv_diff_host_rate
11	num_failed_logins	32	dst_host_count
12	logged_in	33	dst_host_srv_count
13	num_compromised	34	dst_host_same_srv_rate
14	root_shell	35	dst_host_diff_srv_rate
15	su_attempted	36	dst_host_same_src_port_rate
16	num_root	37	dst_host_srv_diff_host_rate
17	num_file_creations	38	dst_host_serror_rate
18	num_shells	39	dst_host_srv_serror_rate
19	num_access_files	40	dst_host_rerror_rate
20	num_outbound_cmds	41	dst_host_srv_rerror_rate
21	is_host_login		

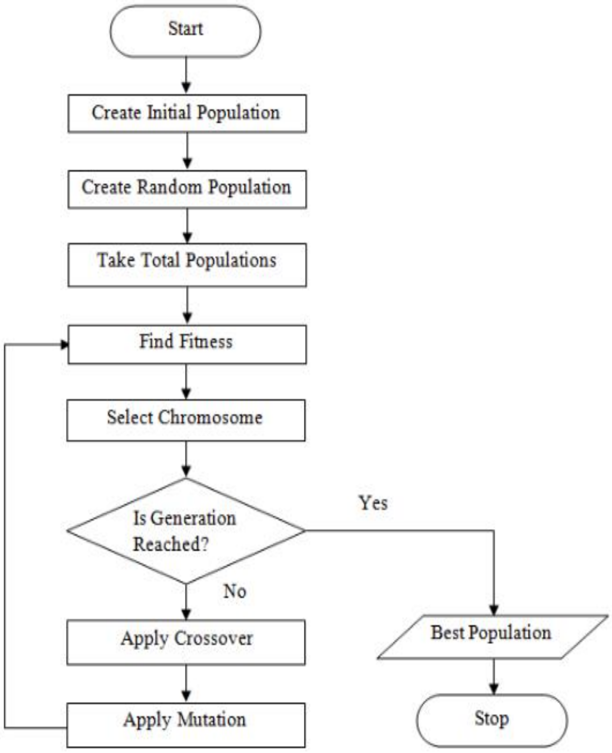


Fig.2. Flow chart of Genetic Algorithm

Rule Generation (Preprocessing)

In data preprocessing basic features are get removed from the dataset. As NSL KDD comprises of 41 distinct features as for packet however from them just a few features will get select. Those are as per the following properties:

a) Duration

b) Protocol

c) Flag

d) Service

e) Source byte

f) Destination byte

g) Class

Initial, an arbitrarily produced population of potential solutions is made. At that point crossover, mutation and choice are connected to every era until a satisfactory solution is found or some time breaking point is surpassed. Crossover is the place two people swap successions of bits to shape two new people. Crossover takes two guidelines and makes new principles by swapping the bits of the old standards. Mutation is the place irregular bits in an individual, or conceivable solution, are arbitrarily changed. The fitness of an individual is indicated by the fitness work, which decides the nature of a specific person.

Capture Data Packet

The packet era may do in different ways i.e. packet can be created from the Summon incite or by utilizing any apparatus, for example, wireshark. In the wake of catching the packets the features are sent towards the IDS then IDS will go to test these packets.

Apply SVM

Support Vector Machine is for the most part utilized for the classification of the given articles. These packets then recognized with the goal that they can be utilized for obviously recognizing the packets about their inclination whether they are strange in nature or consummately so for that reason the classification is finished by the SVM.

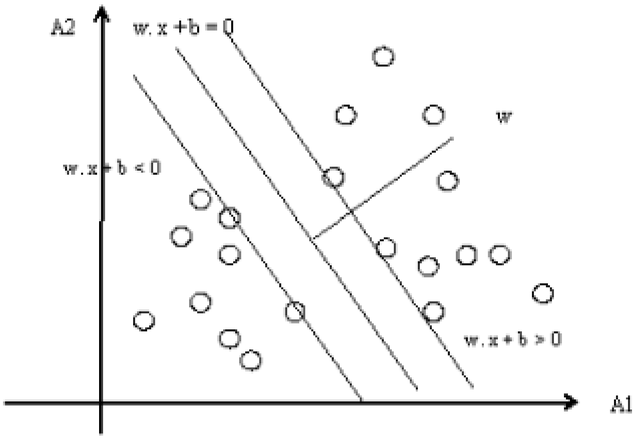


Fig.3. SVM Classification

Algorithm Steps:

Input: A training set with each data point labeled as positive or negative (class labels)

Output: A classifier

1) Begin

2) Randomly select data points from each class.

3) Generate a SVM classifier.

4) While more points to add to training set

```
5) Find support vectors among the selected points;
6) Apply clustering around the support vectors;
7) Add the points in the clusters to the training set;
8) Retrain the SVM classifier using the updated training set;
9) End
```

Apply Ant Colony

This algorithm is chiefly utilized for improving throughput of the given framework. The ant colony grouping takes the contribution from the SVM which has characterized the given packets into various class in light of the way of packets. At that point these packets will be given to the ant colony algorithm where they get coordinated with some predefined design.

```
1) Begin
2) Normalize the data;
3) Let r be the detection rate, initially 0;
4) While r RR do
5) for k = 1, N do
6) SVM training phase;
7) Ant clustering phase;
8) End
9) Construct classifiers;
10) do testing to update r;
```

4. RESULT ANALYSIS

Results are generated and placed in the log files and graphs are generated by taking the values from log files.

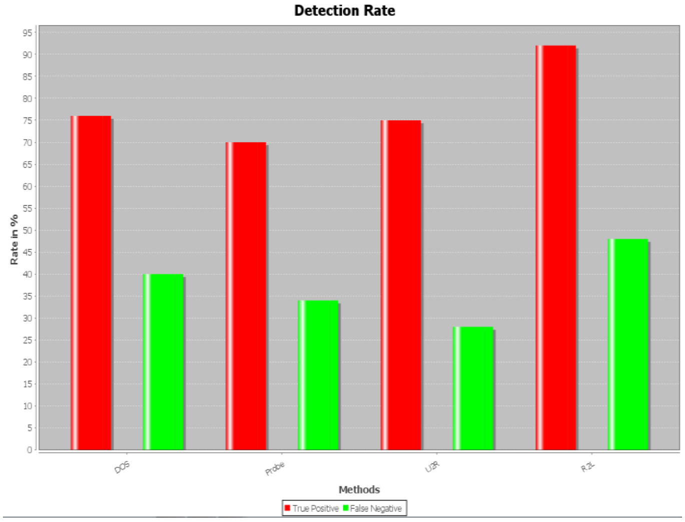


Fig.4. Detection rate graph

These graph shows types of attacks detected. True positive rate is the fraction of intrusions correctly diagnosed (detected), and False negative rate is the fraction of intrusions incorrectly diagnosed (not detected).

Example

70 Intrusion and 100 packets out of which 58 are correct dignose,42 are incorrect then

True Positive = 58/70 = 0.823 *100=82.3%

False Negative = 100-82.3=17.7%

5. CONCLUSIONS AND FUTURE WORK

Decision tree gives better precision for Probe, R2L and U2R classes contrasted with SVM and it gives more awful exactness for DoS class of assaults. For typical class both give a similar execution. There is a little contrast in the exactness for Normal, Probe and DoS classes for decision trees and SVM yet there is a noteworthy distinction for U2R and R2L classes. These two classes have little preparing information contrasted with different classes, so we can reason that decision tree gives great exactness with little preparing informational collections. The outcomes likewise demonstrate that testing time and preparing time of the classifiers are marginally superior to SVM. In addition, decision tree is fit for multi-class classification which is impractical with SVM. Multi-class classification is an exceptionally valuable element for intrusion detection models. The preparation time and testing times are additionally less for decision tree contrasted with the SVM (Support Vector machine).

ACKNOWLEDGEMENT

For everything I achieved, the credit goes to all those who had really helped us to complete this work successfully. Extremely thankful to Prof. Arvind R. Bhagat Patil for guidance and review of this paper also thanks the all faculty members.

REFERENCES

[1] Cuiwei Li,Qin Tu,MaozhengZhao,JunXu,Aidong Men,Amultiscale compressed video saliency detection model based on ant colony optimization, 2015 IEEE/CIC International Conference on Communications in China (ICCC) Year: 2015

[2] Hongxin Liu; JunzhongJi; Cuicui Yang; JiaweiLv; Xiuzhen Zhang,Ant Colony Clustering Approach Combined with Multilevel Framework for Functional Module Detection in Large-Scale PPI Networks , 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)

[3] Abhishek Gupta; Om JeePandey; MahendraShukla; Anjali Dadhich; Anup Ingle; Vishal Ambhore,Intelligent Perpetual Echo Attack Detection on User Datagram Protocol Port 7 Using Ant Colony Optimization , 2014 International Conference on Electronic Systems, Signal Processing and Computing Technologies

[4] Brian C. Williams; Errin W. Fulp, A Biologically Modeled Intrusion Detection System for Mobile Networks , 2010

[5] International Conference on Broadband, Wireless Computing, Communication and Applications Xiaojing Yuan; Zehang Sun; Y. Varol; G. Bebis ,A distributed visual surveillance systemProceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003.

[6] Zohreh Sadat Hosseini; Seyyed Javad Seyyed Mahdavi Chabok; Seyyed Reza Kamel, DOS intrusion attack detectionby using of improved SVR , 2015 International Congress on Technology, Communication and Knowledge (ICTCK)

[7] FaridLawan Bello; KiranRavulakollu; Amrita,Analysis and evaluation of hybrid intrusion detection system models , 2015 International Conference on Computers, Communications, and Systems (ICCCS .

[8] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[9] Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[10] Y.T. Yu, M.F. Lau, "A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions", Journal of Systems and Software, 2005, in press.

[11] Fan Li, Hybrid Neural Network Intrusion Detection System Using Genetic Algorithm 2010 International Conference on Multimedia Technology.

[12] Fatemeh Barani, A hybrid approach for dynamic intrusion detection in ad hoc networks using genetic algorithm and artificial immune system ,2014 Iranian Conference on Intelligent Systems (ICIS).

[13] Jungwon Kim; P. J. Bentley , "Towards an artificial immune system for network intrusion detection: an investigation of dynamic clonal selectionEvolutionary Computation, 2002. CEC '02.Proceedings of the 2002.

[14] Mohammad Saniee Abadeh; Jafar Habibi; Emad Soroush, "Induction of Fuzzy Classification Systems Using Evolutionary ACO-Based AlgorithmsFirst Asia International Conference on Modelling Simulation (AMS'07)