IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 11, November 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

© 2024 IJIRCCE | Volume 12, Issue 11, November 2024| DOI: 10.15680/IJIRCCE.2024.1211082

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Effective & Presentative Modeling of Intrusion Detection using Evolutionary Approach Methodology

Sachin Ahirwar[1], Prof. Sarvesh Site[2]

M. Tech. Scholar, Department of Computer Science and Engineering, All Saints' College of Technology, Bhopal, India[1]

Guide, Department of Computer Science and Engineering, All Saints' College of Technology, Bhopal, India[2]

ABSTRACT: Now days, the fast rising networks proliferation, data transfer rate, and an unpredictable Internet usage have added more anomaly problems. Thus researchers need to develop more reliable, effective, and self-monitoring systems, which sort troubles and can carry out operation devoid of human interaction. By undergoing this kind of attempts, catastrophic failures of susceptible systems can be reduced. Detection stability and detection precision are two key indicators used to evaluate IDS (Intrusion Detection System). In this dissertation we present the comparative experimental study for the intrusion detection and our simulation stats that our proposed method gives better results than the previous techniques. Here we proposed a new model for the intrusion detection in a host based and network based, here we used the evolutionary approach such as genetic algorithm for the proposed methods and compare with the existing technique i.e. classification method. Here we used the MATLAB simulator for the detection of intrusion and the input dataset is k d d c u p 99.

KEYWORDS: Intrusion Detection System, KDDCUP99, MATLAB

I. INTRODUCTION

Intrusion is an unwanted activity in the network and intrusion detection is an important research and development topic with many applications that influencing confidentiality, integrity, availability. In 2014, according to research of Forbes, the most ruthless intrusions include cyber attack stealing personal records of users of eBay, intrusion to Montana Health Department, intrusion effecting P.F. changes customers by stealing their credit and debit card numbers, and finally intrusions affecting evernote and feedly users. It is clear that intrusion detection is so important for a good security policy. There are two main approach for security management these approaches are prevention- based and detection-based.

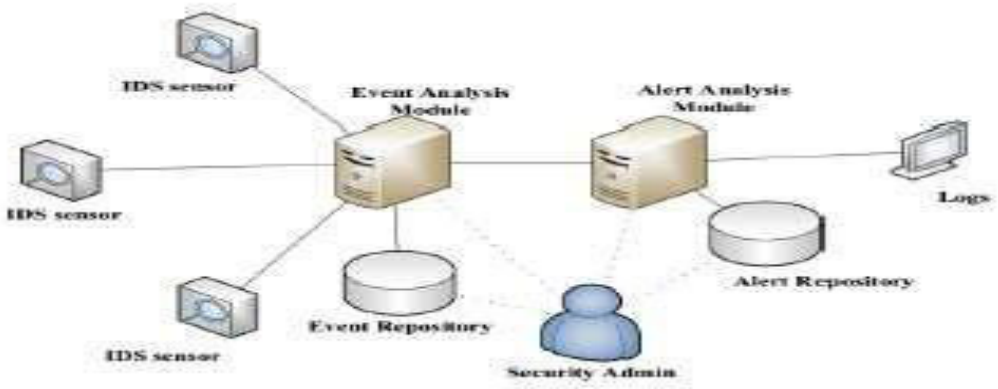

Figure 1: General Architectural diagram of IDS

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

In any security plan, if intrusion prevention (encryption, authorization, and authentication) named as the first line of security is passed by attackers, as a second line of defence, intrusion detection comes into prominence. Intrusion detection provides deterrence for intruder and serves an alarm mechanism for a computer system or a network to manage security plan successfully. An intrusion-detection system (IDS) can be defined as software or hardware tools that monitoring network to detect internal or external cyber attacks. An Intrusion Detection System can observe and investigate system and user activities, recognize patterns of known attacks, identify abnormal network activity. General definition of IDS is about intrusions to network but for WSN it can be added that physical damages to sensor devices. Identifying sensor damage is important in order to serve fault tolerance and reliability [1, 2].

With the high usage of Internet in our day today life, security of network has become the key foundation to all web applications, like online auctions, online retail sales, etc. Detection of Intrusion, attempts to detect the attacks of computer by examining different information records observed in network processes.

This can be considered as one of the significant ways to effectively deal with the problems in network security. An intrusion in the internet can compromise the data security through several internet means. Nowadays, the fast rising networks proliferation, data transfer rate, and an unpredictable Internet usage have added more anomaly problems. Thus researchers need to develop more reliable, effective and self- monitoring systems, which sort troubles and can carry out operation devoid of human interaction. By undergoing this kind of attempts, catastrophic failures of susceptible systems can be reduced. Detection stability and detection precision are two key indicators used to evaluate IDS (Intrusion Detection System).

II. PROPSOED METHODOLOGY

Genetic algorithm is heuristic function; the nature of genetic algorithm is optimal. In this dissertation we used genetic algorithm for optimizations for cluster for classification of data. The genetic algorithm improved the efficiency of cluster with classification model. Now here we discuss the working process of genetic algorithm. Genetic Algorithm (GA), first introduced by John Holland in the early seventies, is the powerful stochastic algorithm based on the principles of natural selection and natural genetics, which has been quite successfully, applied in machine learning and optimization problems. To solve a Problem, a GA maintains a population of individuals (also called strings or chromosomes) and probabilistically modifies the population by some genetic, operators such as selection, crossover and mutation, with the intent of seeking a near optimal solution to the problem. Coding to Strings in GA, each individual in a population is usually coded as coded as a fixed-length binary string.

2.1 Initial Population

The initial process is quite simple. We create a population of individuals, where individual in a population is a binary string with a fixed-length, and every bit of the binary string is initialized randomly.

2.2 Evaluation

In each generation for which the GA is run, each individual in the population is evaluated against the unknown environment. The fitness values are associated with the values of objective function.

2.3 Genetic Operators

Genetic operators drive the evolutionary process of a population in GA, after the Darwinian principle of survival of the fittest and naturally occurring genetic operations. The most widely used genetic operators are reproduction, crossover and mutation. To perform genetic operators, one must select individuals in the population to be operated on .The selection strategy is chiefly based on the fitness level of the individuals actually presented in the population. There are many different selection strategies based on fitness. The most popular is the fitness proportionate selection. After a new population is formed by selection process, some members of the new populations undergo transformations by means of genetic operators to form new solutions (a recombination step). Because of intuitive similarities, we only employ during the recombination phase of the GA three basic operators: reproduction, crossover and mutation, which are controlled by the parameter pr, pc and pm(reproduction probability, crossover probability and Mutation probability), respectively. Let us illustrate these three genetic operators. As an individual is selected, reproduction operators only copy it from the current population into the new population (i.e., the new generation) without alternation. The crossover operator starts with two selected individuals and then the crossover point (an integer between 1 and L-1,

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

where L is the length of strings) is selected randomly. Assuming the two parental individuals are x1 and x2, and the crossover point is 5 (L=20). If

X1= (01001|101100001000101)
X2= (11010|011100000010000)

Then the two resulting offspring are

X'1= (01001|011100000010000)
X'2= (11010|101100001000101)

The third genetic operator, mutation, introduces random changes in structures in the population, and it may occasionally have beneficial results: escaping from a local optimum. In our GA, mutation is just to negate every bit of the strings, i.e., changes a 1 to 0 and vice versa, with probability pm.

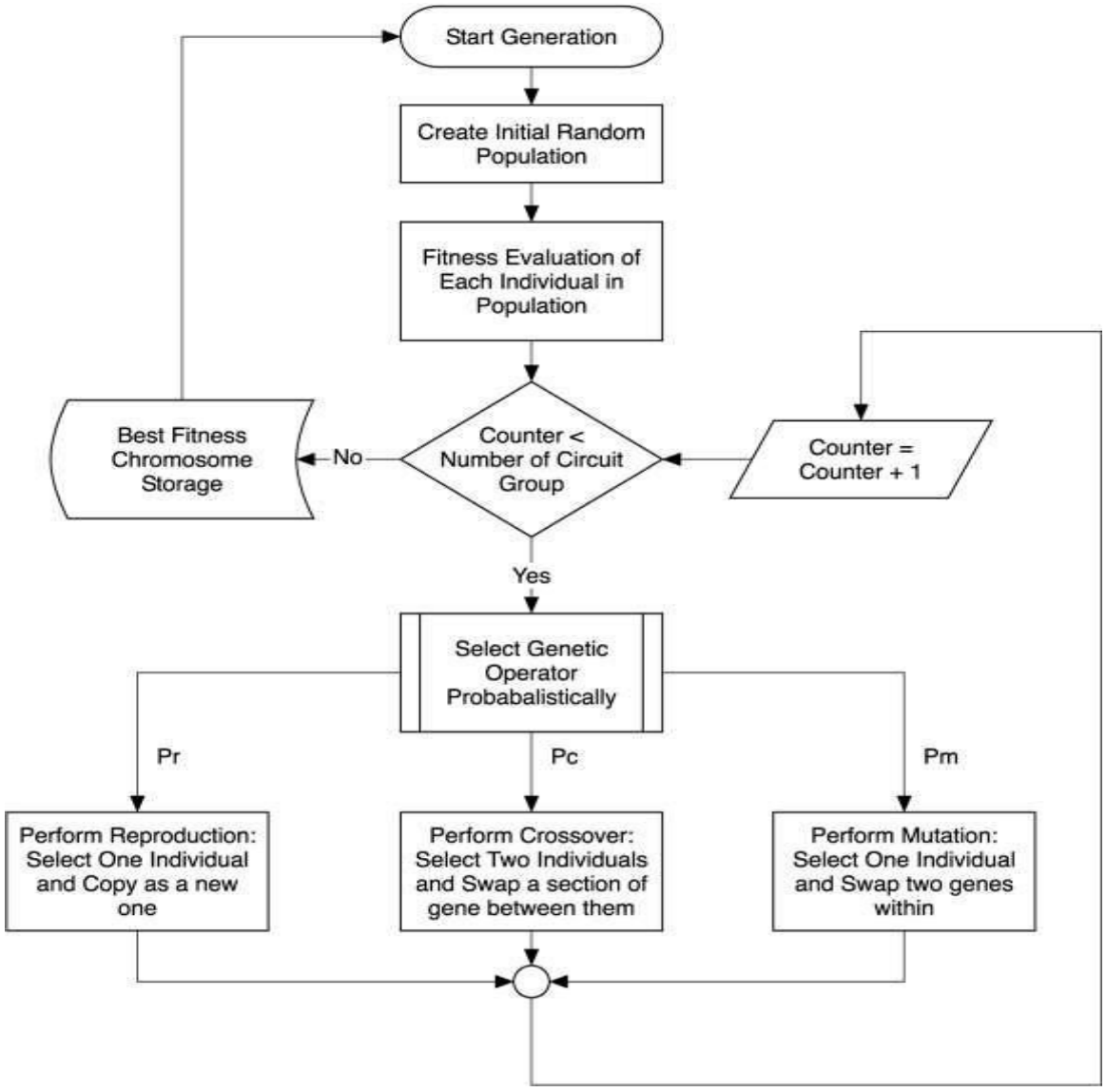

Figure 2: Shows that block diagram of working principle of genetic algorithm

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. Generate random population of *n* chromosomes (suitable solutions for the problem)
2. Evaluate the fitness *f(x)* of each chromosome *x* in the population
3. Create a new population by repeating following steps until the new population is complete

a. Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected)

b. With a crossover probability cross over the parents to form new offspring (children). If no crossover was performed, offspring is the exact copy of parents.

c. With a mutation probability mutate new offspring at each locus (position in chromosome).

d. Place new offspring in the new population

4. Use new generated population for a further run of the algorithm
5. If the end condition is satisfied, **stop**, and return the best solution in current population
6. Go to step **2.**

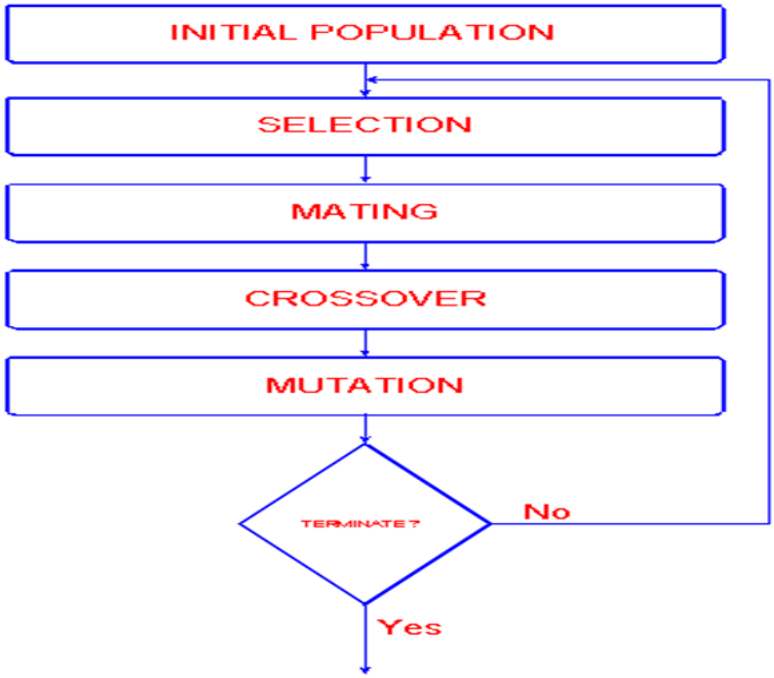

Figure 3: Flow graph for the genetic algorithm procedure.

Our system can be divided into two main phases: the pre-calculation phase and the detection phase. Listing 1 depicts major steps in pre-calculation phase, where a set of chromosome is created using training data. This chromosome set will be used in the next phase for the purpose of comparison.

Listing 1. Major steps in pre-calculation
Algorithm: Initialize chromosomes for comparison Input : Network audit data (for training)
Output: A set of chromosomes

1. Range = 0.125
2. For each training data
3. If it has neighboring chromosome within Range
4. Merge it with the nearest chromosome
5. Else
6. Create new chromosome with it
7. End if
8. End for

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

III. SIMULATION PARAMETER

In this section we discuss about the performance evaluation parameters for the proposed method with compare with the other existing method, all parameters which is based on some alarm signal generated by the our proposed intrusion detection model and compute the results on that basis.

Precision: measures the proportion of predicted positives/negatives which are actually positive/negative.

$$Precision = \frac{TP}{TP+FP} * 100 \ldots\ldots\ldots\ldots\ldots (1)$$

Recall: It is the proportion of actual positives/negatives which are predicted positive/negative.

$$Recall = \frac{TP}{TP+FN} * 100 \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Accuracy: It is the proportion of the total number of prediction that were correct or it is the percentage of correctly classified instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} * 100 \ldots\ldots\ldots (3)$$

IV. SIMUALTION RESULT

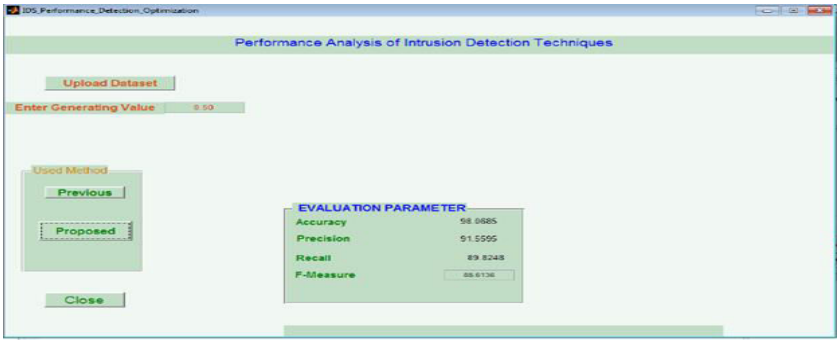

Figure 4: Shows that the intrusion data classification, when the number of generating value is 0.5 and the method is Proposed

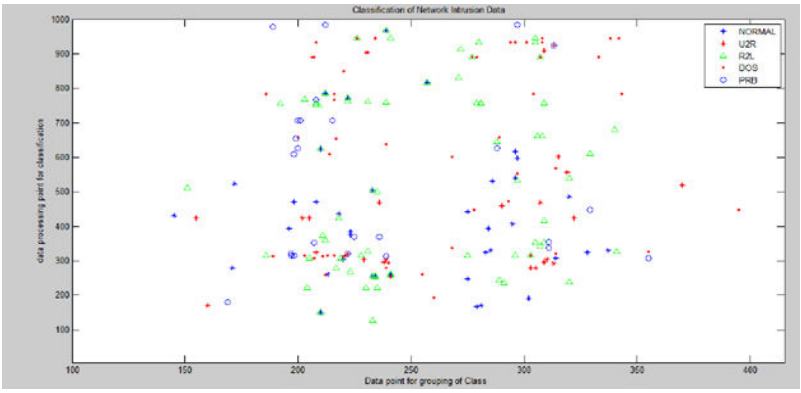

Figure 5: Shows that the intrusion data classification, when the number of generating value is 0.5 and the method is proposed

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

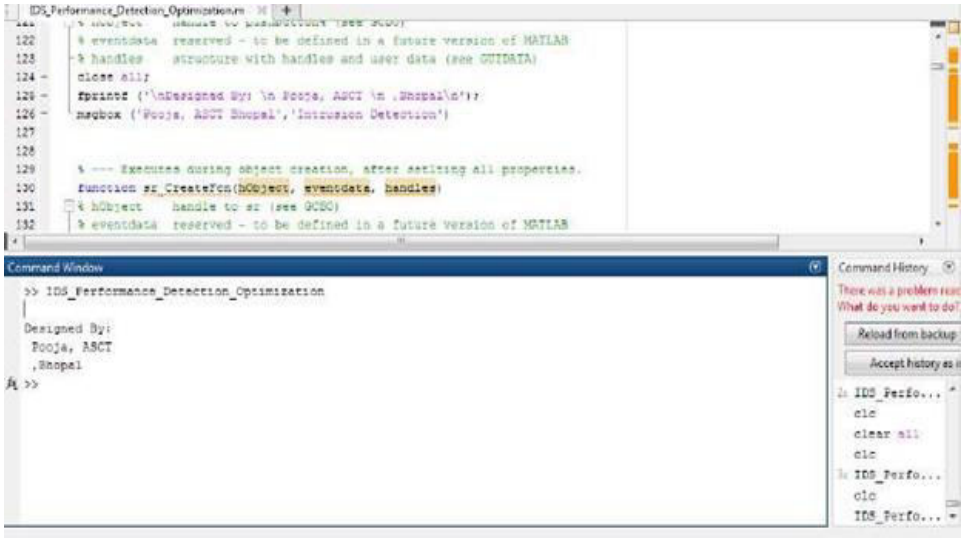

Figure 6: Shows that exit environment window for the simulation process.

Table I: Shows that the performance parameter evaluation of given input value such as 0.9 for the previous and proposed method

INPUT VALUE	METHOD	ACCURACY	PRECISION	RECALL
0.9	Previous (SC4ID Algorithm)	95.27	88.91	84.38
	Proposed (GENETIC Algorithm)	96.48	89.64	88.70

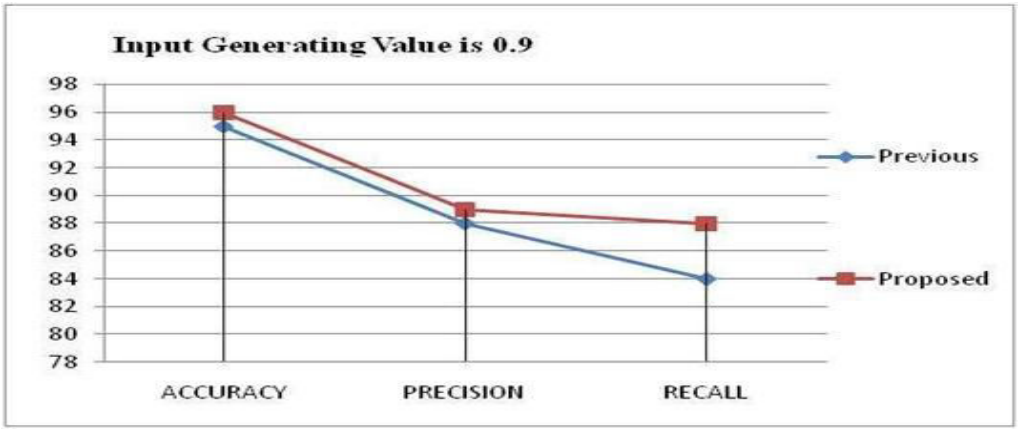

Figure 7: This figure shows that the comparative experimental study for the previous and proposed method the generating input value is 0.9

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. CONCLUSION

The balance between detection rate and false positive rate become more challenging when normal activity and anomalous activity are not static. The activity on the network can change and the IDS must be aware of this change and adapt accordingly. If not, the ability of the IDS to provide accurate and reliable results is greatly diminished. Therefore, an IDS must adapt to different environments, which potentially bring different activity and behavior unseen by the IDS. In this dissertation we present the comparative performance evaluation for the intrusion detection using the classification and genetic algorithm method, Our simulated result shows that the proposed method gives better results in terms of accuracy, precision and recall than previous classification method.

REFERENCES

[1] Pierre-Francois Marteau, "Sequence Covering for Efficient Host-Based Intrusion Detection", IEEE Transactions on Information Forensics and Security, Vol. 14, No. 4, APRIL 2019, pp 944-1006.
[2] Rashidah Funke Olanrewaju, Burhan Ul Islam Khan, Athaur Rahman Najeeb, Ku Nor Afiza Ku Zahir and Sabahat Hussain, "Snort-Based Smart and Swift Intrusion Detection System", Indian Journal of Science and Technology, VOL 11(4), DOI: 10.17485/ijst/2018/v11i4/120917, January 2018.
[3] Ashima Chawla, Brian Lee, Sheila Fallon, and Paul Jacob, "Host Based Intrusion Detection System with Combined CNN/RNN Model", Springer Nature Switzerland August 2019, pp 149-158.
[4] Md. Zahangir Alom, Venkata Ramesh Bontupalli, and Tarek M. Taha, "Intrusion Detection using Deep Belief Networks", IEEE 2015, pp 339-344.
[5] Hebatallah Mostafa Anwer, Mohamed, Farouk, Ayman Abdel-Hamid, "A Framework for Efficient Network Anomaly Intrusion Detection with Features Selection", IEEE 2018, pp 157-162.
[6] Nutan Farah Haq, Musharrat Rafni, Abdur Rahman Onik, "Application of Machine Learning Approaches in Intrusion Detection System: A Survey", (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.3, 2015, pp 9-18.
[7] Rana Aamir Raza Ashfaq , Xi-Zhao Wang , Joshua Zhexue Huang , Haider Abbas , Yu-Lin He , "Fuzziness based semi-supervised learning approach for intrusion detection system", Elsevier 2016, pp 484-497.
[8] Anna L. Buczak and Erhan Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection", IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 18, NO. 2, SECOND QUARTER 2016, pp 1153-1176.
[9] JABEZ J, Dr.B.MUTHUKUMAR, "Intrusion Detection System (IDS): Anomaly Detection using Outlier Detection Approach", Procedia Computer Science 48 2015, pp 338 – 346.
[10] M Firoj kabir,Sven Hartman, "Cyber Security:Challenges An efficient Intrusion Detection System Design",IEEE 2018, pp 19-24.
[11] Poonam Sinai Kenkre, Anusha Pai, and Louella Colaco, "Real Time Intrusion Detection and Prevention System", Springer International Publishing Switzerland 2015, pp 405-411.
[12] Pierre-Francois Marteau, "Sequence Covering for Efficient Host-Based Intrusion Detection", JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. X, FEBRUARY 2018, pp 1-14.
[13] G. Yedukondalu, J. Anand Chandulal and M. Srinivasa Rao, "Host-Based Intrusion Detection System Using File Signature Technique", Springer Nature Singapore Pte Ltd. 2017, pp 225-232.
[14] Okan CAN, Ozgur Koray SAHINGOZ, "A Survey of Intrusion Detection Systems in Wireless Sensor Networks",IEEE 2015, pp 1-6.
[15] Shijoe Jose, D.Malathi, Bharath Reddy, Dorathi Jayaseeli, "A Survey on Anomaly Based Host Intrusion Detection System", NCMTA 2018, pp 1-11.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details