Volume 3, Issue 11, November 2013 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Genetic Algorithm based Detection of Malicious Activities in the Network

Karamjeet Kaur*, Er. Navdeep Singh

Department of Computer Engineering,
University College of Engineering,
Punjabi University, Patiala, Punjab, India

Abstract- Now a day, it is very important to maintain a high level security to ensure safe and trusted communication of information between various organizations. But secured data communication over internet and any other network is always under threat of intrusions and misuses. So intrusion detection system has become a needful component in terms of computer and network security. Intrusion detection is one of the important issues in network security. Anomaly detection is one of the most important intrusion detection techniques which detect the unknown attacks. In this paper we propose a genetic algorithm for intrusion detection. The genetic algorithm generates the optimal solution for the problem. Using genetic algorithm the probability of obtaining the optimal solution is increased. The performance of proposed method is compared with multistage anomaly detection technique. In this proposed work we uses the unweight dataset for the intrusion detection.

Keywords- Intrusion detection, anomaly detection, network based intrusion detection, genetic algorithm.

I. INTRODUCTION

Intrusion detection system is a system that detects the abnormal or malicious activities in the internet. Intrusion detection system can be categorized into two parts. The host based intrusion detection system can detects the attacks on the one system but network based intrusion detection system can detects the intrusion from the multiple hosts. The two types of intrusion detection techniques are used for intrusion detection. The misuse detection technique can detects the known virus signature but anomaly detection technique can detects the known patters and unknown patterns from the network. Recently various data mining approaches have been proposed for intrusion detection such as classification, clustering, association mining and frequency episode. Genetic algorithm is an optimization algorithm which is used for to find the optimal solution. The process of genetic algorithm starts with randomly selected population of chromosomes. These chromosomes are used for the problem to be solved. According to the problem, different positions of each chromosome are encoded as bits, characters or numbers. These positions are referred to as genes and are changed during evolution. The set of chromosomes are called as population. An evaluation function is used to calculate the goodness of every chromosome. The two basic operators are crossover and mutation are used to generate the rules and mutation operator can replace the lowest fitness value of chromosome with highest fitness value of chromosome. Finally the best individual is packed out as final result when the optimization criterion is met.

In this paper, we propose a detection of malicious activities in the network using the genetic algorithm. In this work unweight dataset is used for anomaly detection. This work follows network based intrusion detection using anomaly detection technique.

II. LITERATURE SURVEY

Yao et al. describes the fuzzy intrusion detection system in detail. Fuzzy logic is a self adaptive algorithm, which detects the intrusions with threshold value. This paper also describes the functional components of intrusion detection.

B. Benet et al. integrates the fuzzy logic with genetic algorithm. The genetic algorithm is used for optimization of fuzzy rules. In this paper the rules are generated using the fuzzy logic and genetic algorithm chooses the better or good rules for intrusion detection.

Lixin Wang was point out the weakness of the previous methods such as statistical analysis and rule based system in intrusion detection. This paper describes that how we can detect unknown attacks and avoid malicious hiding intrusion, artificial neural network for anomaly detection was introduced. He said that ANNs may be the most suitable technology for anomaly intrusion detection.

Dusan Stepanovic and Slobodan Bojanic proved intrusion detection with good accuracy using genetic algorithm. Genetic algorithm (GA) field is one of the up-coming fields in computer security, especially in intrusion detection systems (IDS). GA operates on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to the solution of the problem that GA is trying to solve.

III. DESIGN PROCESS

A. Database

A database is an organized collection of data. The unweight dataset is used for intrusion detection. We can perform the intrusion detection on unweight dataset.

B. Selection method

The selection method is used for selecting the input from the input data for intrusion detection. The numbers of rules are also used with selection method. When we select the input data from the database then we apply the rules on data and check that whether this data is normal or abnormal. So selection method denied the intrusion communications and allows the communication between normal connections.

C. Fitness function

The fitness function is used for to check the goodness of rules.

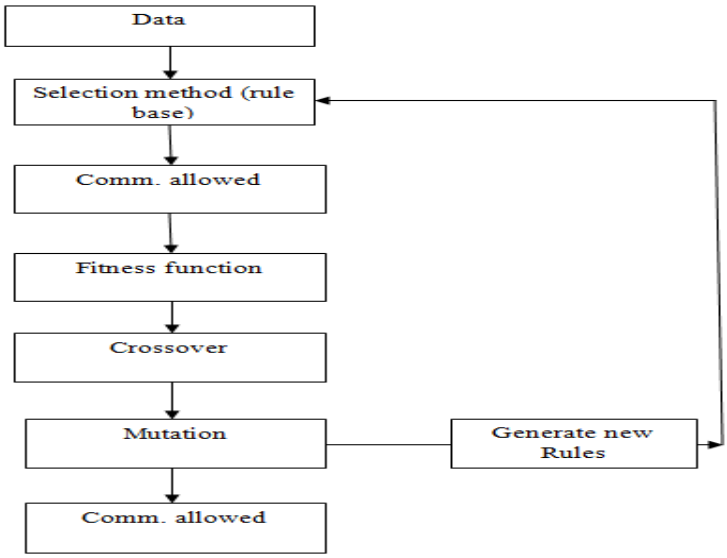

Fig 1- Design process

D. Crossover operator

Crossover is a genetic operator that combines two chromosomes to produce two new chromosomes. The reason of crossover is that the new chromosome may be better than both of the parent if it takes the best characteristics from each of the parent.

E. Mutation operator

Mutation is a genetic operator that alters one or more gene value in a chromosome from its initial state. In result, new gene value being added to gene pool. With these new gene values, the genetic algorithm may produce better solution than previously. After generating a new rules then we apply these rules on the data. If the intrusion is found then we deny that communication nodes otherwise communication will be allowed between the nodes.

IV. EXPERIMENTAL RESULTS

In the dissertation work we have implemented a Genetic algorithm for intrusion detection in data. From the qualitative and quantitative results, we observe that genetic algorithm works properly and detects the intrusions efficiently.

A. Qualitative results

The qualitative results are those which can be observed but not measured.

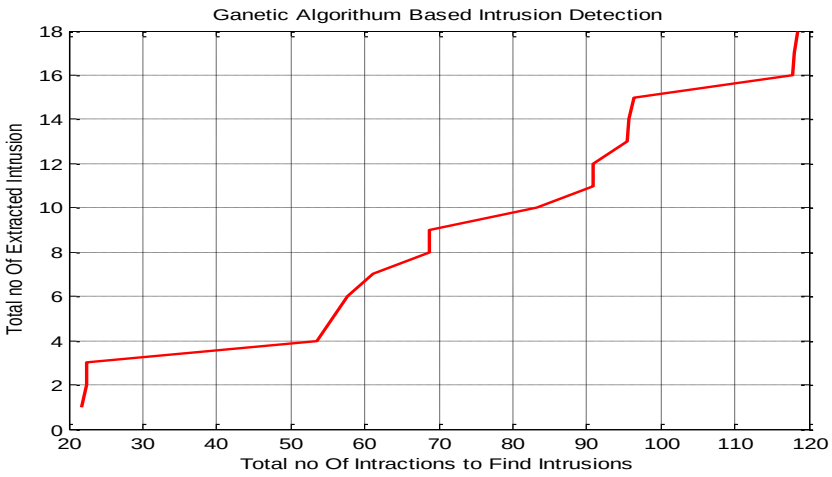

Fig 2- Intrusion detection with population size is 50.

The qualitative results in figure 2 shows the intrusion detection in y-axis and total number of interactions is shown in x-axis. If we select the population size 50 then 18 intrusions are detected, shown in figure.

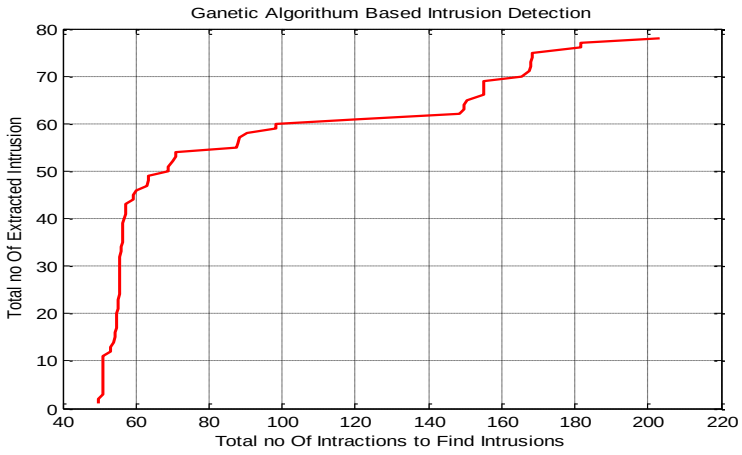

Fig 3- Intrusion detection with population size is 100

In this result (fig 3) we increase the population size or input size up to 100. So the figure shows that by increasing the population size we detect the 78 intrusions with number of interactions.

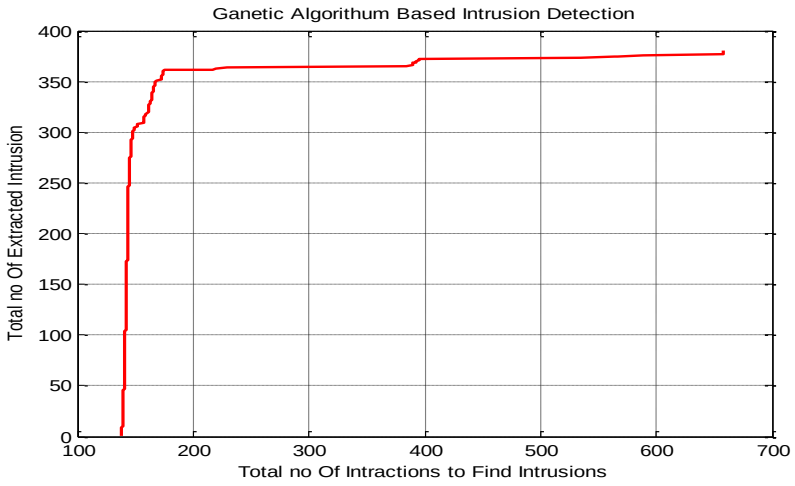

Fig 4- Intrusion detection with population size 150

From figure 4 we observe that the number of intrusions is more as compared to the previous result. In this result we increase the population size or input size up to 150. If new type of data is coming then genetic algorithm automatically generates the new types of rules. So the figure shows that by increasing the population size we detect the 360 intrusions with number of interactions.

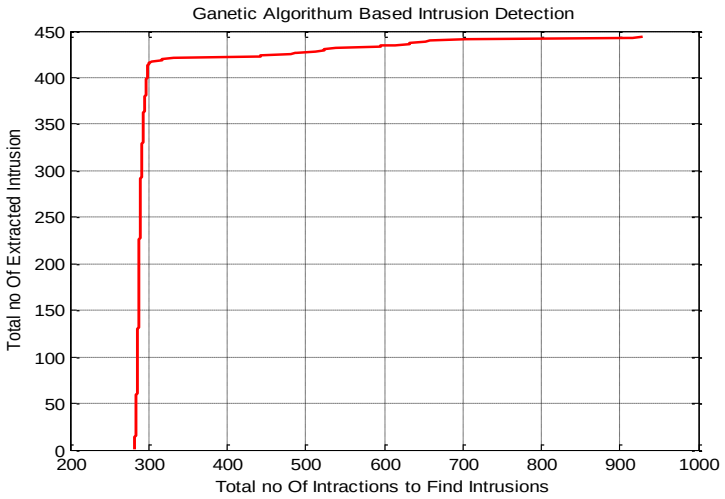

Fig 5- Intrusion detection with 200 population size.

In this result fig 5 we increase the population size or input size up to 200 and we detect the 440 intrusions with number of interactions as shown in results.

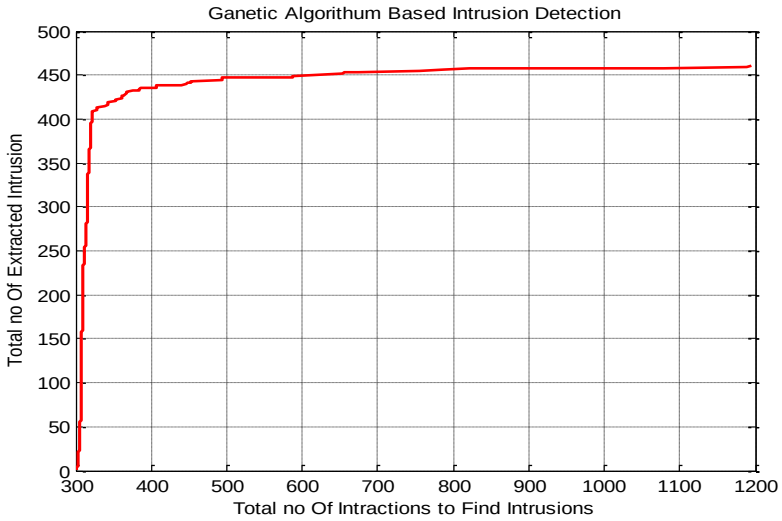

Fig 6- Intrusion detection using 350 population size.

From the fig 6 we observe if we increase the population size or input size up to 350 then we detects the 460 intrusions.

B. Quantitative results

The result which indicates some numbers or some dimensions, known as quantitative results.

TABLE 1
DETECTION OF SUSPECTED INTERACTIONS.

P_size	Allow time	Total infections	Interactions	Total time
50	2437	130	2500	0.74
100	5413	3 11	5500	2.32
150	8311	1176	8400	15.13
200	12426	1611	12600	27.17
350	23748	1588	24000	44.07
440	30429	1790	30800	58.68

From the above quantitative results we observe that if we increase the population size then infections or intrusions are also increased or may decreased. The P_size is the population size which we increased. The allow time indicates the start time of every stage. The infections indicates the number of detected intrusions.

1) Comparisons with previous results

TABLE 2
ACTUAL NUMBER OF INTRUSIONS AND DETECTED INTRUSION

Population size	Actual number of intrusions	Intrusions detected by proposed method	Percentage
50	181	130	68.78
100	380	311	81.84
150	1424	1176	82.58
200	1806	1611	89.2
350	1851	1588	85.75
440	1895	1790	94.45
Average percentage			**83.77**

In table 2, the actual number of intrusions and detected intrusions are presented. The percentage of all detected intrusions is also calculated. The actual number of intrusions is those intrusions which are occurred in the dataset. These are predefined intrusions which are present in the dataset. The average percentage is 83.77 percent.

TABLE 3
COMPARISON WITH PREVIOUS RESULTS.

IDS	Detection Rate (Detected attacks/detectable attacks)
Expert-1	50.3% (85/169)
Expert-2	46.8% (81/173)
Dmine	40.2% (41/102)
Forensics	55.6% (15/27)
NETAD	71.4% (132/185)
Parallel system	60.8% (104/171)
Multi-Stage Network anomaly detection	68.4% (117/171)
Proposed system	87.64% (6606/7537)

Table 3 shows the comparison between proposed technique and existing techniques. This table shows the intrusion detection system and detection rate. In the previous research the detection rate is 68.4% and we achieve the detection rate is 87.64%.

V. CONCLUSION

The aim of this research work is to implement the post-data analysis engine using intrusion detection system for analysis using genetic algorithm. Genetic algorithm based post data analysis IDS can lead to the future of research in the area of post network data analysis which can helpful to find out the intrusions and infectious content in the network using gathered data. This system will act as the black box, which will gives us all of the information about the infectious happenings in the network. This algorithmic technique is specifically designed to represent the results in the graphical form using the genetic algorithm. This research may lead the network analysis and research field into the innovative post analysis field.

FUTURE SCOPE

The algorithmic results are able to detect the intrusions and infectious content in the previously happened communications by using post network communication data. This algorithm will lead the research into the field of post network data analysis to find out the possible intrusions which may not get detected in the real communications. These results and data can act as the black box in the network. All of the network communications or interactions data may be stored in some sort of server or workstation in the network. Data collected in that particular server or workstation would be kept safe for post network data analysis.

REFERENCES

[1] Reddy Kesavulu, *"A study of intrusion detection in data mining"*, 2011.
[2] Yao and Zhao, *"A study on fuzzy intrusion detection*", 2007.
[3] Wei li, "*Using genetic algorithm for network intrusion detection*" 2003.
[4] Liao, Tian shengfeng, "*Network forensics based on fuzzy logic and expert system*", science direct:1881-1892.
[5] Simranjeet and Neeta, "*Soft computing in intrusion detection*", 2010.
[6] John and ali, "Network *intrusion detection using an improved competitive learning neural network*", 2008.
[7] K. Ilgun and A. Kemmerer, "*State transition analysis: A rule-based intrusiondetection approach*", IEEE Transaction on Software Engineering 21(3): 181-99 (1995).
[8] Kandeeban and S rajesh, "*Integrated intrusion detection system using soft computing*", IJNS: 87-92, 2010.
[9] Zorana Bankovic, stepnvic, "*Improved network security using genetic algorithm approach*" , science direct: 438-451 , 2007.
[10] Abadeh saniee, licas.c "*Intrusion detection using a fuzzy genetic based learning algorithm*", science direct: 414-428, 2007
[11] Toosi adel and Kahani, " *A new approach on intrusion detection using genetic algorithm and neural network*" science direct : 2007.
[12] Anup goyal and chetan, "*A genetic algorithm network intrusion detection system*", 2008.
[13] Islam and ahmed ," *Fuzzy grid based intrusion detection in neural network*". 2012.
[14] Ben, Roja and Paramesvri, "*Intrusion detection using fuzzy genetic algorithm*":IJARCSSE, 2012.
[15] Luo, susan M bridge, "*Mining fuzzy association rules and fuzzy frequency episodes for intrusion detection*", 2000.
[16] Kshirsagar and M. Tidke, "*Intrusion detection system using genetic algorithm and data mining*": 2231-2232, 2012.
[17] Ramesh babu, "*Intrusion detection using data mining along fuzzy logic and genetic algorithms*", IJCSNS:vol-8, 2008.
[18] Disha Sharma, "*An intrusion detection using clustering techniques*" IEEE: 2011.

[19] Varun Chandola Anomaly Detection for Symbolic Sequences and Time Series Data, PhD. Dissertation. Computer Science Department, University of Minnesota,2009.
[20] Ojugo, A.O Eboka, " *Genetic algorithm rule based intrusion detection system*" , CIS journal:1182-1190, 2012.
[21] Morgan Kaufman, "*An Immunity-based Technique to Characterize Intrusions in Computer Network"* Transactions on Computation, Vol. 6(3), pp. 281-291. GECCO, pp.1081-1088.
[22] R.A. Kemmerer and G. Vigna "*Intrusion Detection: A Brief History and Overview*", IEEE Computer, Vol.2 ,2002.
[23] Przemyslaw Kazienko and Piotr Dorosz "*Intrusion Detection Systems (IDS) Part 2 Classificationmethods;techniques"* web white paper, 2004.
[24] Tarek S. Sobh and Wael M. Mostafa, "A *cooperative immunological approach for detecting network anomaly*", Applied Soft Computing, Elsevier, Vol. 11(1), pp. 1275-1283, 2011.
[25] P. Garcia Teodorro, J. Diaz-Verdejo, G. Marcia- Fernandez, E. Vazquez, "*Anomaly-based network intrusion detection: Techniques, systems and challenges*", Computers and Security" Vol.28(1-2), pp.18-28, 2009.
[26] A. Aziz, Salama, Sanaa El-Ola "*Artificial Immune System Inspired Intrusion Detection System Using Algorithm*" 2012.
[27] Kumar, Upendra "*An efficient intrusion detection based on decision tree classifier using feature reduction"* 2012.
[28] Anand, swaraj, bgarat "*genetic algorithm based intrusion detection*", IJCST: vol-3,2012.
[29] Botha, von solmn, "*Utilising fuzzy logic and trend analysis for effective intrusion detection*", 2003.

