PARAMETERS FOR DERIVING FITNESS FUNCTION OF GENETIC ALGORITHM IN INTRUSION DETECTION

Gnanaprasanambikai.L
Research Scholar, C.M.S College of Science and Commerce,
Coimbatore, India.
gnanaambikai@gmail.com.

Dr. Nagarajan Munnusamy
Associate Professor, C.M.S College of Science and Commerce,
Coimbatore, India.
mnaagarajan@gmail.com.

Abstract Securing network system from several new attacks is a challenge in recent years. Hence intrusion detection is becoming one of the key security components in network security. Several soft computing approaches are applied for applying intrusion detection in recent years, Genetic Algorithm is a powerful, robust optimization soft computing approach which is suitable for intrusion problem. The success of Genetic algorithm process depends on Fitness function which is used in genetic algorithm. This paper suggests various existing fitness functions. The paper also proposes parameters for deriving new fitness function.

***Keywords*:** Intrusion Detection, Genetic Algorithm, Fitness Function.

1. INTRODUCTION

Rapid growth of Internet, computer systems are facing number of security threats. To preserve the integrity, confidentiality and available of data, in computer many efforts have been made like encryption, firewall, anti-virus software etc. With this Intrusion detection is also added as a compliant to detect malicious behaviors when occurs [1]. Genetic Algorithm applied on Intrusion Detection increases high detection rate.

In this paper an analysis on various fitness function of genetic algorithm used in intrusion detection is done. Parameters for constructing a fitness function is suggested by considering the drawbacks of various fitness function.

2. INTRUSION DETECTION SYSTEMS

Intrusion Detection (ID) is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of intrusions.

A. Based on location

Intrusion Detection System (IDS) is classified in to two categories: Host IDS (HIDS) and Network IDS (NIDS). HIDS run on individual hosts or devices on the network. A HIDS monitors the inbound and outbound packets from the device only and will alert the user or administrator for any suspicious activity detected. NIDS are placed at a strategic point or points within the network to monitor traffic to and from all devices on the network [2]

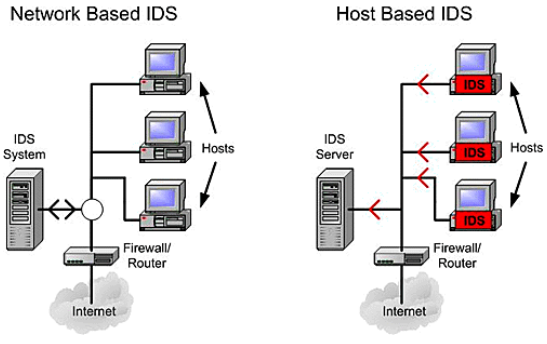

Figure 1[16] Host IDS Vs Network IDS

B. **Based on Detection**

IDS are classified into two categories based on detected method: Signature ID and Anomaly ID. Signature Intrusion Detection is also known as Misuse Intrusion Detection. This Detection will monitor packets on the network and compare them against a database of signatures or attributes from known malicious threats. Anomaly Intrusion Detection measures deviations from normal baseline [1].

Table 1. Misuse versus Anomaly

	Advantages	**Dis-Advantages**
Misuse Intrusion Detection	Detect attacks accurately	New attacks are not Detected
Anomaly Intrusion Detection	Novel attacks are Detected	Produce more False positives

3. GENETIC ALGORITHM

Genetic Algorithm is an evolutionary and soft computing approach to support intrusion detection. Genetic algorithm is a problem solving method that uses genetics as its model of problem solving. It's a search technique to find approximate solutions to optimization and search problems. GA handles a population of possible solutions of optimization problem. Each solution is represented through a chromosome, which is just an abstract representation [3].

Genetic Algorithm work begins with a set of solutions called initial populations. The solutions are evaluated by fitness function and sequence of operations namely selection, crossover, mutation and replacement are applied. The base for all operations is fitness function [4].The process is repeated until the termination condition becomes true. The termination condition is a convergence criteria which may be maximum number of generations, elapsed time, no improvement in the fitness function of the chromosomes [3].

Pseudo code
for population of best fit rules(local optimum) instead of global optimum.
Begin
Initialize population with random chromosomes
Evaluate each chromosome for fitness
Repeat until termination condition is satisfied do
1. Select chromosomes
2. Recombine gene of chromosomes
3. Mutate resulting offspring
4. Select chromosomes for the next generation
End

4. IMPLEMENTING OF GA IN IDS

Genetic Algorithm is used in number of ways in Intrusion Detection Systems. Genetic Algorithm features of Optimization [5], Parallelism [6], Adaptability [6], Robustness [3], and Evaluation function [3] are suitable for Intrusion Detection Problem. Genetic Algorithm is an adaptive and heuristic technique based on natural genetics. The rule in intrusion detection mimics the chromosome data structure of genetic algorithm. Attributes of the rule mimics the genes of a chromosome in a genetic algorithm [5]. The Genetic algorithm supports Intrusion detection system searching for population of best fit rules (local optimum) instead of global optimum.

5. TERMINOLOGIES USED IN INTRUSION DETECTION SYSTEM

A rule in intrusion detection system consists of two parts one is condition (antecedent) and other is action (consequent). To evaluate intrusion detection following terminologies are followed [1]

True Negative - If condition and action both are true means the rule classifies normal data as normal.
False Positive - If condition is false and action is true means the rule classifies normal data as intrusion.
False Negative - If condition is true and action is false means the rule classifies intrusion data as normal
True Positive - If condition is false and action is false means the rule classifies intrusion data as intrusion

6. LITERATURE SURVEY ON FITNESS FUNCTION

Fitness function is used to check the fitness of a rule in the training stage of genetic algorithm. If the rule (Chromosome in genetic algorithm) is fit, it used for reproduction. When Genetic algorithm applied for intrusion detection, number of fitness function is proposed. Each has its own way of evaluating a chromosome (rule in intrusion detection). The following are most commonly used fitness functions of GA in Intrusion Detection System.

[7] used GA for intrusion detection system, he calculated fitness function by the following four equations,

$$Outcome = \sum_{i=1}^{57} Matched * Weight(i)$$

$$\Delta = |\text{Outcome} - \text{Suspiciouslevel}|$$

$$Penalty = (\Delta * Ranking)/100$$

$$fitness = 1 - penalty$$

[8][9][10] used support-confidence framework as fitness function. The fitness function of rule is given by support –confidence framework

$$Support = |A \text{ and } B|/|N|$$

$$Confidence = |A \text{ and } B|/|A|$$

$$Fitness = w1 * support + w2 * confidence$$

Where N is the total number of connections in audit data, |A| stands for the number of network connection matching the condition A, and |A and B| is the number of network connections that matches the rule if A then B. The weights w1 and w2 are to control the balance between the two terms and have the default values of w1=0.2 and w2=0.8. [8] used the fittest rules detecting novel attacks on networks to detect novel or unknown networks in genetic programming implementation. [9] used fittest rules clearly to classify the types of intrusions and about 97% of attacks were detected correctly.

[11][12][13] has presented a genetic algorithm to identify the harmful attack type of connections. [11] Fitness function was derived for misuse intrusion detection. The fitness function given by the formula

$$\textbf{Fitness} = \frac{\boldsymbol{a}}{\boldsymbol{A}} - \frac{\boldsymbol{b}}{\boldsymbol{B}}$$

Where a is the number of correctly detected attacks, A is the total number of attacks, b is the number of normal connections, B is the total number of normal connections in the training set.[12] used three features in rule to have high detection rate 95.72% and low false 4.27% rate. [13] used rule classify normal and abnormal behaviors using eighteen features set with high detection rate of 99.87% and .003% low false positive rate.

[14] used Reward penalty based fitness function for genetic algorithm for intrusion detection. The Fitness Function given by the formula

$$\text{Fitness} = 2 + \frac{\text{AB-A}}{\text{AB+A}} + \frac{\text{AB}}{\text{X}} - \frac{\text{A}}{\text{Y}}$$

[15] used following fitness function to measure a chromosome strength.

$$\textbf{Fitness} = \mathbf{f(x)/f(sum)}$$

Where f(x) is the fitness of individual x and f(sum) is the entire fitness of all individuals. [15] produced the detection rate of 91.025% .

7. PARAMETERS FOR FITNESS FUNCTION DERIVATION

In Genetic Algorithm process the base for all operations is fitness function. When Genetic algorithm applied for intrusion detection, various fitness functions are derived with different detection rates. The paper proposes following drawbacks of existing fitness function as the parameters for deriving new fitness function

- [7] The fitness function uses more equations which is difficult and time consuming.
- [8][9][10] The fitness function is difficult in process and time consuming.
- [11][12][13]The fitness function is not precise and used in both anomaly and misuse intrusion detection which mislead terminologies of intrusion detection
- [14]The Fitness function uses more number of terms for its evaluation which is difficult and confusing
- [15] The fitness function is not precise where individual fitness is not mentioned.
- [11] The fitness function is not precise in setting threshold value.
- [8] The fitness function is not precise is setting weight value.

8. CONCLUSION AND FUTURE WORK

The paper suggests the importance of fitness function of Genetic Algorithm in intrusion detection. The drawbacks of existing fitness function and considerations for deriving new fitness function. As a future work fitness function is derived based on the parameters mentioned.

9. REFERENCES

[1] K.G.Srinivasa and N.Pramod, "gNIDS: rule-based network intrusion detection systems using genetic algorithms", International Journal of Intelligent Systems Technologies and Applications, vol 11, Nos 3/4, pp 252-266, 2012.

[2] Brijendra Singh, "Network Security and Management", PHI Learning Pvt Ltd, Second edition 2009.

[3] S.N.Sivanandam, S.N.Deepa, "Introduction to Genetic Algorithms", ISBN 978-3-540-73189-Springer.

[4] P.G.Majeed, S.Kumar, "Genetic Algorithms in Intrusion Detection Systems: A Survey, International Journal of Innovation and Applied Studies, ISSN 2028-9324 Vol. 5 No. 3, pp. 233-240, Mar. 2014.

[5] http://www.myreaders.info/08_Fundamentals_of_Genetic_Algorithms.pdf.

[6] K.K.Prasad and S.Borah, "Use of Genetic Algorithms in Intrusion Detection Systems:Analysis","International Journal of Applied Research and studies", ISSN: 2278-9480 Volume 2, Issue 8, Aug – 2013.

[7] Wei. Li, "A Genetic Algorithm Approach to Network Intrusion Detection", SANS Institute, USA, 2004.

[8] W.Lu and I.Traore, "Detecting New forms of Network Intrusion Detection using Genetic Programming", Computational Intelligence, Blackwell Publishing, Malden, vol.20, pp.475-494,2004.

[9] R.H.Gong , M.Zulkernine and P. Abolmaesumi, "A software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection", Proceedings of the 6th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks, , Towson, Maryland, USA,pp:246-253, May 23-25, 2005.

[10] A.A.Ojugo,A.O.Eboka,O.E.Okonto,R.E.Yoro,F.O.Aghware, "Genetic Algorithm Rule-Based Intrusion Detection System", "Journal of Emerging Trends in Computing and Information Science", Vol.3, No.8, August 2012.

[11] A.Goyal and C.Kumar, "Genetic Algorithm based Network Intrusion Detection System", from: http://www.cs.northernwestern.edu/~ago210/ganids/GANIDS.pdf

[12] V.M.Hashemi, Z.Muda and W.Yassin, "Improving Intrusion Detection Using Genetic Algorithm", Information Technology Journal 12(11): pp 2167-2173, 2013.

[13] B.Abdullah, I.Abd-alghafar, G.I. Salama, A.Abd-alhafez, "Performance Evaluation of a Genetic Algorithm Based Approach to Network Intrusion Detection", 13th International Conference on Aerospace Sciences & Aviation Technology, ASAT-13, 2009, pp: 1-17.

[14] F.Alabsi and R.Naoum, "Fitness Function for Genetic Algorithm used in Intrusion Detection System," International Journal of Applied Science and Technology, Vol.2, No.4, PP.129-134, April 2012.

[15] Priya U. Kadam, P. P. Jadhav, "An effective rule generation for Intrusion Detection System using Genetics Algorithm", International Journal of Science, Engineering and Technology Research, Volume 2, Isue 10, October 2013.

[16] http://www.informit.com/articles/article.aspx?p=29601