# Absolute Fitness function of Genetic Algorithm For Traffic Anomaly Intrusion Detection System

[1]L.Gnanaprasanambikai, [2]Dr. Nagarajan Munusamy

[1]*Assistant Professor, Nehru Arts and Science College, Coimbatore*

[2]*HOD, Associate Professor, K.S.G College of Arts and Science,Coimbatore*

[1]*gnanaambikai@gmail.com*

[2] *mnaagarajan@gmail.com*

***Abstract-***Intrusion Detection System a imperative tool used in network security which detects normal and abnormal packets in the network. Genetic Algorithm is a evolutionary algorithm which exploits the tolerance of imprecision, partial truth, uncertainty and approximates to achieve robustness. Genetic Algorithm is applied to Intrusion Detection Problems of rule fitness. The base of Genetic Algorithm operations is fitness function. In this paper we propose a fitness function for Anomaly Intrusion Detection and reason for proposing by specifying the lacking of existing fitness function.

***Keywords-*** *Intrusion Detection, Misuse, Anomaly, Genetic Algorithm, Fitness Function*

## I. INTRODUCTION

Security of a network is very important in an organization. Organizations implement many methods in securing their network system and data namely firewall, antivirus software's, cryptographic methods etc. In these securing mechanisms, Intrusion detection is one such method for securing network and host in a network for preserving data in a host. Intrusion detection is a monitoring tool which informs network administrators about abnormal packet transfer through a network. Based on detection, Intrusion Detection is classified into two categories namely Misuse intrusion detection and Anomaly intrusion detection. Misuse Intrusion Detection validates the packet in network for some known attacks. Anomaly Intrusion Detection validates the packet for unknown (or novel) attacks. Each method has its own merits and de-merits. Misuse Intrusion Detection give accurate results but lacks in finding new attacks as there is no profile for unknown attacks already stored. Anomaly Intrusion Detection finds new attacks but lack in accurate result as it suspects all packets [8]. In this article we discuss the issues of Anomaly Intrusion Detection and the method of problem solving.

## II. GENETIC ALGORITHM

Genetic Algorithm is an adaptive heuristic search. Genetic Algorithm is an evolutionary algorithm used in optimization and search process. Genetic algorithm is a method of problem solving that uncovers estimated results to optimization and search problems. GA handles a collection of feasible solutions for optimization problem. Each solution is represented through a chromosome, which is conceptual representation. Genetic Algorithm work begins with a collection 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. These operations are repetitive until the termination condition becomes true. The termination condition is also known as convergence criterion. It may be maximum number of generations, elapsed time, and no improvement in the fitness function of the chromosomes [9].

## III. APPLYING GENETIC ALGORITHM IN INTRUSION DETECTION

In anomaly intrusion detection, to validate the packet, a rule is needed. The accuracy of packet detection depends on the fitness of rule. The rule plays a vital role and it is the solution. The reason for choosing Genetic Algorithm for intrusion detection is characteristics of Genetic Algorithm are suitable for solving Intrusion Detection Problem. The characteristic of Genetic Algorithm includes Optimization, Parallelism, Adaptability [5], Robustness, and Evaluation function are suitable for Intrusion Detection Problem. Genetic Algorithm is a problem solving method which finds the fittest rules from the population of the rules. The Genetic Algorithm fitness function determines the quality of the rule. The best global rules are selected by applying the Genetic Algorithm operations. The below figure shows the basic structure of Genetic Algorithm function. The Genetic Algorithm operations cannot be executed without the fitness evaluation. The Genetic Algorithm depends on the fitness function.

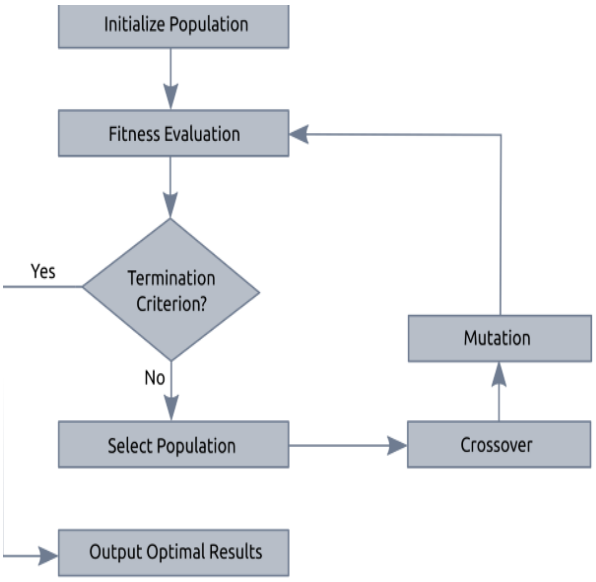


**Figure 1: Basic Structure of Genetic Algorithm work**

## IV. TERMS USED IN INTRUSION DETECTION

A rule in intrusion detection system evaluates a packet as normal and abnormal. It consists of two parts condition and action part. The rule of intrusion detection is of the form:

If condition (antecedent) then action (consequent)

The condition part of a rule is a collection of attributes. The rules in intrusion detection imitate the chromosomes of genetic algorithm and elements of the rule imitate the genes of chromosome in genetic algorithm.

To evaluate intrusion detection following terminologies are followed [8]

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

## V. FITNESS FUNCTION – RELATED WORKS

Fitness function is used to check the fitness of a rule in the Genetic algorithm process. If the rule is fit, it is used for real time implementation. When Genetic algorithm applied for intrusion detection, number of fitness function is proposed. Each has its own way of evaluation.

[11] Use GA for anomaly intrusion detection system, the fitness function is calculated by the following four equations,

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

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

$$\text{Penalty} = \frac{V * \text{Ranking}}{100}$$

$$\text{Fitness} = 1 - \text{Penality}$$

[12][7][2] Use support-confidence framework as fitness function. The fitness function of rule is given by support –confidence framework

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

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

$$\text{Fitness} = w1 * \text{support} + w2 * \text{confidence}$$

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. [7] used the fitness for anomaly intrusion detection. [2] used fittest rules clearly to classify the types of intrusions and about 97% of attacks were detected correctly.

[1][10][3] presented a genetic algorithm to identify the harmful attack type of connections. [1] Fitness function was derived for misuse intrusion detection. The fitness function given by the formula

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

Where a is the number of correctly detected attacks that is True Positive, A is the total number of attacks, b is the number of normal connections that are falsely identified as attacks that is False Positive, B is the total number of normal connections in the training set. [10] used three features in rule to have high detection rate 95.72% and low false 4.27% rate for anomaly intrusion detection. [3] Used rule for anomaly intrusion detection with eighteen features set with high detection rate of 99.87% and .003% low false positive rate.

[4] Used Reward penalty based fitness function for genetic algorithm for intrusion detection where reward is given to chromosome strength and penalty is given to chromosome weakness.

The Fitness Function given by the formula

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

The fitness function is used for Misuse Intrusion Detection. The Fitness function contains large number of terms and equations.

[6] Used following fitness function to measure chromosome strength.

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

Where f(x) is the fitness of individual x and f(sum) is the entire fitness of all individuals. [6] produced the detection rate of 91.025% and individual fitness formula is not specified.

## VI. PROPOSED FITNESS FUNCTION

Section 5 suggests commonly used fitness function in Genetic Algorithm operations for intrusion detection system. The suggested fitness functions are not precise whether to be used in misuse intrusion detection or anomaly intrusion detection system since both systems have both their own strength and weakness. The existing fitness functions are not clear in setting threshold values whether to use constrain optimization or un-constrained optimization methods.

We propose a Fitness function for the rule in Anomaly Intrusion Detection which consider all terms of Intrusion Detection.

$$\text{True fitness} = \left(\frac{TP}{TP+FN}\right) + \left(\frac{TN}{TN+FP}\right) \quad \text{------Eqn(1)}$$

$$\text{False Fitness} = \left(\frac{FP}{TN+FP}\right) + \left(\frac{FN}{TP+FN}\right) \quad \text{------Eqn(2)}$$

$$\text{Fitness} = \text{Eqn}\,(1) - \text{Eqn}\,(2)$$

The fitness value is high means the survival of the rule is high and low means the survival of the rule is low. Survival of the rule high means the rule undergoes genetic algorithm operations for longer duration, which results fittest rule for real time implementation. In the proposed fitness function False Negative is considered for evaluation.[1][10][3] Used only false positive for evaluation. A good Fitness function should considers all terms of intrusion detection system with all false values are to be reduced and all true values are to be increased. The proposed fitness function solves the anomaly intrusion detection weakness by using all terms of intrusion detection system.

## VII. PROPOSED FITNESS FUNCTION RESULTS

We implement the proposed fitness function in Traffic anomaly intrusion detection using NSL KDD dataset. NSLKDD is the de facto dataset for anomaly intrusion detection. For traffic anomaly network intrusion detection, we generated the rules with the feature set of Multiple connection derivative features, which specifies the count of connections to a particular destination IP address and port. We implemented the proposed fitness function using MATLABR2014 tool.

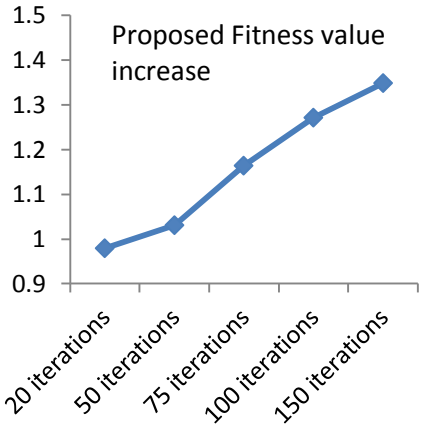


**Figure 2. Graph Showing Proposed Fitness value at Several Iterations**

The figure above shows the Global fitness value of several iterations using proposed function. The fitness value increases as iteration increases favoring increase in true values and reduce in false values.

## VIII. CONCLUSION AND FUTURE WORK

In this paper the Genetic Algorithm role in Intrusion detection systems, Fitness function role in Genetic Algorithm operations and related works are discussed. The drawback of existing Fitness function is mentioned and a new fitness function is proposed for anomaly intrusion detection. The reason for the proposed fitness function also discussed. The Future work is to implement the proposed fitness function to give a comparison with existing results and proposed results to validate the proposed fitness function.

## REFERENCES

[1] AnupGoyal and C.Kumar (2008).Genetic Algorithm based Network Intrusion Detection System. http://www.cs.northernwestern.edu/~ago210/ganids/GANIDS.pdf

[2] A.A.Ojugo, A.O.Eboka, O.E.Okonto, R.E.Yoro, F.O.Aghware (2012). Genetic Algorithm Rule-Based Intrusion Detection System. Journal of Emerging Trends in Computing and Information Science. 3(8).

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

[4] F.Alabsi and R.Naoum (2012). Fitness Function for Genetic Algorithm used in Intrusion Detection System. International Journal of Applied Science and Technology. 2(4). 129-134

[5] K.K.Prasad and S.Borah (2013). Use of Genetic Algorithms in Intrusion Detection Systems: Analysis. International Journal of Applied Research and studies 2(8). 2278-9480.

[6] Priya U. Kadam, P. P. Jadhav (2013). An effective rule generation for Intrusion Detection System using Genetics Algorithm. International Journal of Science, Engineering and Technology Research. 2(10).

[7] R.H.Gong, M.Zulkernine and P. Abolmaesumi (2005). 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. 246-253.

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

[9] Sivanandam.S, .N, Deepa. S.N (2008). Introduction to Genetic Algorithms. ISBN 978-3-540-73189-Springer.

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

[11] Wei. Li (2004). A Genetic Algorithm Approach to Network Intrusion Detection. SANS Institute, USA.

[12] W.Lu and I.Traore (2004). Detecting new forms of Network Intrusion Detection using Genetic Programming. .Computational Intelligence. Blackwell Publishing, Malden. 475-494.