International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 4 335 - 339

An Approach for Intrusion Detection by using "Genetic Algorithm

Mr.Devendra Vaidya
Student, M.Tech, Fourth Semester, CSE. Department
RGCER, hingna
Nagpur,Maharashtra, India
E-mail: devvaidya123@gmail.com

Prof. Aswini Yerlekar
Professor ,CSE dept.RGCER,Hingna ,Nagpur,
Maharashtra,India
E-mail:ashwini.yerlekar@gmail.com

***Abstract*—**As computer attacks are becoming more and more difficult to identify the need for better and more efficient intrusion detection systems increases. The main problem with current intrusion detection systems is high rate of false alarms. Distributed Denial of Service (DDoS) attacks are large-scale cooperative attacks launched from a large number of compromised hosts called Zombies, Which are a major threat to Internet services. Therefore, keeping this problem in view here presents various significant areas where genetic algorithm techniques seem to be a strong technique for detecting and preventing DDoS attack. Our purpose of this work is to examine how to integrate multiple intrusion detection sensors in order to minimize the number of incorrect alarms.so a brief introduction to the parameters and evolution process of a GA will be provided by this process and how to implement it in real IDS.

Keywords*: Distributed Denial of Service attack, Genetic Algorithm, Zombies, intruders ,intrusion detection*

*****

I. INTRODUCTION

The primary issue with current interruption recognition frameworks is the high rate of false alerts activated off by aggressors. Powerful method for ensuring the system against malignant assaults is the issue in both territory of exploration and the PC system overseeing experts. Enhanced checking of vindictive assaults will require combination of numerous observing frameworks. In our present venture we are dissecting potential advantages of circulated multi sensor frameworks for interruption discovery.

The principal issue is the way to coordinate information from numerous sensors, and the second is the manner by which to distinguish most vital information gave by various sensors. We are creating arrangement of expository and scientific models to utilize potential advantages of different sensors for diminishing false cautions.

The reason for this anticipate is to examine execution of model multi sensor based interruption recognition framework. Today, the quantity of assaults against huge PC frameworks or systems is developing at a fast pace. One of the real dangers to digital security is Distributed Denial-of-Service (DDoS) assault. In DDOS the casualty system element(s) are barraged with high volume of invented assaulting bundles began from countless. The point of the assault is to over-burden the casualty and render it unequipped for performing ordinary exchanges [Kim.et.al. 2004]. Hereditary calculation methodology can be embraced as a beyond any doubt shot weapon to these assaults. The late quick advancement in Genetic Algorithm has made accessible wide assortment of calculations, drawn from the fields of insights, example acknowledgment, machine learning, and database.

The fundamental issue with current interruption location frameworks is high rate of false alerts activated off by aggressors. Powerful securing the system against malevolent assaults remains issue in both examination and the PC system overseeing experts. Enhanced observing of malignant assaults will require coordination of different checking frameworks. A progression of expository and numerical models are utilized to get potential advantages of various sensors for decreasing false cautions. Today, the quantity of assaults against vast PC frameworks or systems is developing at a fast pace.

One of the real dangers to digital security is Distributed Denial-of-Service (DDoS) [3, 5] assault. In which the casualty system element(s) are besieged with high volume of imaginary assaulting bundles started from an extensive number of Zombies. The point of the assault is to over-burden the casualty and render it unequipped for performing ordinary exchanges. To ensure system servers, system switches [4, 8] and customer has from turning into the handlers, Zombies and casualties of appropriated disavowal of-administration (DDoS) assaults. Hereditary calculation methodology can be received as a beyond any doubt shot weapon to these assaults. The focal topic of this paper is to investigate parameters and advancement process[6] of Genetic Algorithm which distinguishes pernicious bundle on the system and at last obstructs the particular IP addresses.

Hereditary calculation is a transformative calculation which is useful for pursuit and enhancement reason. They join the idea of Darwin's hypothesis of survival. Numerous analysts have presented the utilization of GA in interruption Detection and reported high achievement rates. We have utilized GA based way to deal with find and recognize the malignant parcels and IP addresses on the system. The principle explanation for selecting GA for this assignment is because of innate developmental treatment in the calculation which permits us to characterize our own particular wellness capacity in view of which just those individuals or guidelines are chosen that fulfill our wellness rule.

With this approach in mind we have designed GA based system and implemented fitness function on the processes of GA. The aim is to get high prediction rate and minimum false positive rates on network traffic captured by the system. The training of the system is carried out on the predefined rules while other resting is done on the real time data sets generated by the firewall system. The results generated after successful execution of the algorithm are thus justifying the choice, performance and applicability of genetic algorithm into the Intrusion Detection System.

II. LITERATURE SURVEY

The Intrusion Detection System has undergone rapid changes and is using new evolved techniques to generate better results. There are several approaches for solving intrusion detection problems assizes on the network.

Wei Li [6] has describe GA based IDS with a methodology of applying genetic algorithm into network intrusion detection techniques. This implementation of genetic algorithm is unique as it considers both temporal and spatial information of DARPA data set Rule Set Rule Base Network Sniffer GA network connections during the encoding of the problem; therefore, it should be more helpful for identification of network anomalous behaviors.

B. Uppalaiah, T. Bharat et al. [7] presents the Genetic Algorithm for the Intrusion detection system for detecting DoS, R2L, U2R, Probe from DD99CUP data set. The architecture of the system along with implementation of the software for the proposed technique is also discussed. The time to get thorough with three features to describe the data will be reduced with a combination of Genetic Algorithm based IDSs. there system is more flexible for usage in different application areas with proper attack taxonomy. Genetic Algorithm detects the intrusion while correlation techniques identify the features of the network connections .The results shows that we have specified set of rules and high Dos, R2L, U2R, Probe attack detect rate. In Optimizing the parameters present in the algorithm reduces the training time.

Srinivasa K G, SaumyaChandra et al.[8] presents IGIDS, where the genetic algorithm is used for pruning best individuals in the rule set database. The process makes the decision faster as the search space of the resulting rule set is much compact when compared to the original data set. This makes IDS faster and intelligent.

Anup Goyal and Chetan Kumar [9] has presented a machine learning approach known as Genetic Algorithm (GA), to identify such harmful/attack type of connections. The algorithm takes into consideration different features in network connections such as type of protocol, network service on the destination and status of the connection to generate a classification rule set. Each rule in rule set identifies a particular attack type. For this experiment, they implemented a GA and trained it on the KDD Cup 99 data set to generate a rule set that can be applied to the IDS to identify and classify different types of attack connections.

Brian E. Lavender [10] proposed the integration of genetic algorithms (GA) into SNORT to enhance SNORT at performing Network Intrusion Detection (NID).

Shaik Akbar et al. [11] presents an algorithm which identifies damaging/attack type connections called Genetic Algorithm. The algorithm considers different features by protocol type, duration, src_bytes in generating a rule set. The Genetic Algorithm is trained on the KDDCUP99 Data Set in order to generate a collection of rules which applied on Intrusion Detection System identifies different types of attacks.

III. GENETIC ALGORITHM

1. INTRODUCTION:

Hereditary calculations are a branch of developmental calculations [8] utilized as a part of hunt and enhancement methods. The three overwhelming elements of a hereditary calculation i.e., determination, hybrid and transformation relate to the organic process: The survival of the fittest (As appeared in Figure 1).

In a hereditary calculation, there is a populace of strings (called chromosomes or the genotype of the genome), which encode and indent arrangements (called people, animals, or phenotypes).[10] Traditionally, arrangements are spoken to in twofold as series of 0s and 1s, however different encodings are likewise conceivable. The advancement as a rule begins from a populace of arbitrarily produced people and develops over eras. In every era, the wellness of each person in the populace is assessed, various people are stochastically chosen from the present populace (in light of their wellness), and adjusted (recombined and perhaps arbitrarily changed) to frame another populace. The new populace is then utilized as a part of the following cycle of the calculation. Generally, the calculation ends when either a most extreme number of people are there in an era, or an attractive wellness level has been gone after the populace. On the off chance that the calculation has ended because of a most extreme number of people, an attractive arrangement might have been come to.

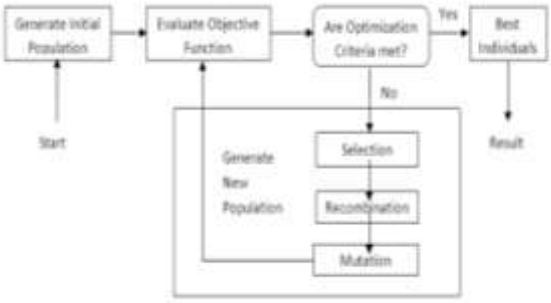

Figure 1: Structure of a Genetic Algorithm

2. GENETIC ALGORITHM PROCESS

GA advances the number of inhabitants in chromosomes (people) as the procedure of characteristic selection.[12] It generate(s) new chromosome(s) (posterity) amid its procedure. GA process utilizes an arrangement of hereditary administrators (determination administrator, hybrid administrator and change administrator), and assess chromosome utilizing the wellness function.GA comprises of populace of chromosomes that imitated over arrangement of eras as indicated by their wellness in a domain. Chromosomes that are most fit are well on the way to survive, mate, and bear children.GA end the procedure by characterize altered maximal number of eras or as the accomplishment of a satisfactory wellness level, or if there are no upgrades in the populace for some settled eras, or for some other reason. The standard GA procedures is appeared

in figure 2. It contains different steps which incorporates: encoding chromosomes, creating introductory populace, wellness capacity assessment, then applying one of the administrators. The procedure will stop when we get the best people.

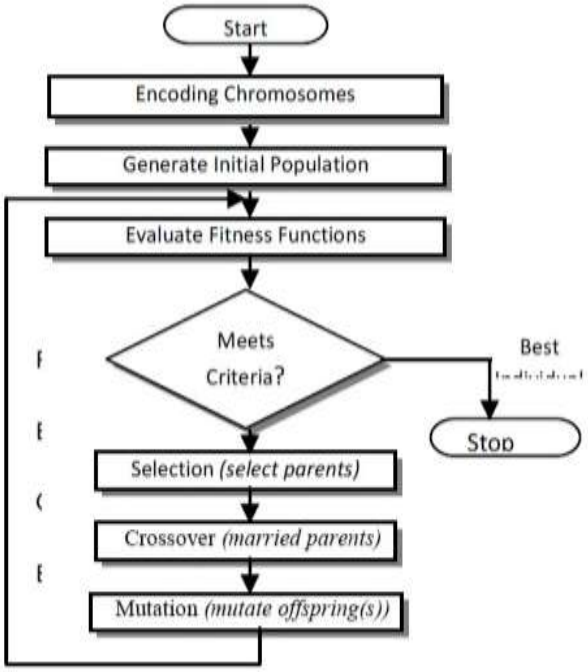

Figure 2: Genetic algorithm Process

3. GENETIC ALGORITHM OPERATORS.

- Encoding of the Chromosomes.

In the GA process it is important to represent the data into some of the encoding formats. There are many encoding methods to represent data string for GA's further process. Like, binary encoding and real valued encoding.

- Applying fitness function.

Fitness function (or objective function) [12] defines the problem constraints; it measures the performance of all chromosomes in the population.

- Selection operator.

Determines which chromosome(s) [12] from the population will be chosen for recombination; depends on the fitness of the chromosome. The selected chromosomes are called parents. Such selection methods are: fitness-proportion selection, roulette-wheel selection, stochastic universal sampling, local selection and rank selection.

- Crossover operator.

The parents' chromosomes are recombined by one of the crossover methods. It produce one or more new chromosome(s) called offspring(s). Such methods are: Single Point Crossover, Multipoint Crossover, Uniform Crossover and Arithmetic Crossover.

- Mutation operator.

New genetic material could be introduced into the new population through mutation process. [12] This will increase the diversity in the population. For each offspring mutation randomly alters some gene(s). Some encoding schemas: binary encoding and real-number encoding.

IV. SYSTEM OVERVIEW

The proposed system overview is shown in figure 3 which starts from capturing firewall entries i.e. firewall data sets and then initial filtering is done on the basis of rule defined by the system. This précised data is then input to the GA based algorithm which generates the best individuals.

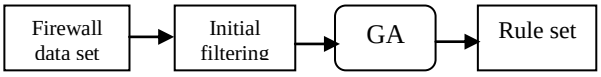

Figure 3: architecture of Genetic Algorithm

The detail proposed architecture is shown in figure 4. It starts from initial population generation from pfirewall.log file generated by the firewall system. The packets are the filtered out on the basis of rules. Then the précised data packets go through several steps namely selection, crossover and mutation operation. These processes gets generate best individuals. The generated individuals are the verified by the fitness function to generate the population for next generation.

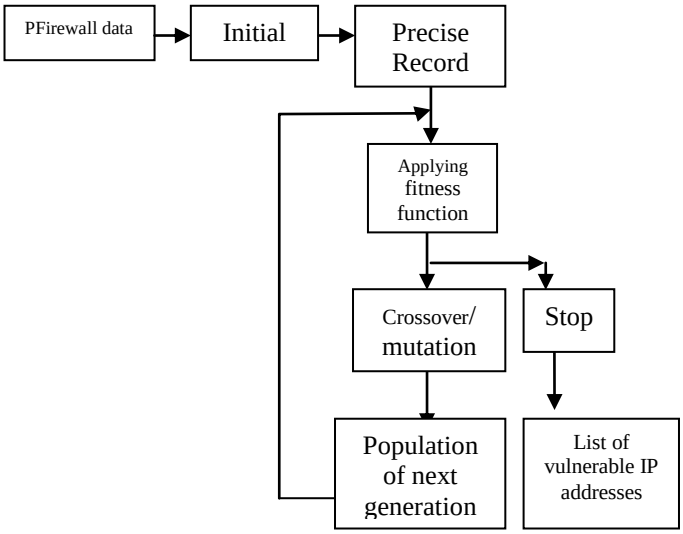

Figure 4: Detailed system architecture for GA-RIDS.

IV. EXPERIMENTAL SET UP

OBJECTIVE:

The IP locations are created on an extensive scale. In Intrusion Detection method the IP addresses making annoyance are should be recognized. The Genetic calculation utilized is to recognize the gatecrashers and the last rundown will be sent to firewall. Consequently the system will be shielded from assaults and will expand the execution of the system

ISSN: 2321-8169
335 - 339

The point of this paper is to distinguish the gatecrashers and create the rundown of IP locations with its event check. The Genetic calculation will distinguish the frail and solid interlopers and produce the last rundown and would piece them on firewall. The slightest access rights given to such clients would expand the execution of the system.

TOOLS

For this experiment we have used java as the frontend software which code overall GA operator and there evolution process. The training data is stored into the wamp server which is used as the backend to the system. For this experiment we used windows based Dell computer with dual core processor system having 120 GB hard disk space and 1 GB RAM to execute the computer program.

- GENETIC OPERATORS AND PARAMETERS

SETTING TYPE	VALUES
Encoding Scheme	Binary Encoding
Population size	100
Evolution generation	50
Selection	Fitness-proportion
Crossover	One point
Mutation	Real number
Generations	50
Fitness function	Weight * pkt_size

Table 1: parameter setting for GA algorithm

V. RESULT

From the above experiment, we have able to create a rule base that could successfully categories harmful and harmless connection types. We have shown the resultant figures below by applying 100 connection entries respectively to the proposed system. After that we were able to get around 96% of accuracy to classify the connections types.

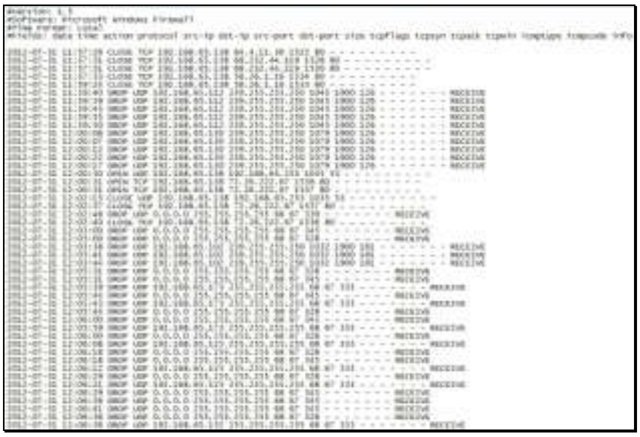

Figure 5: pfirewall.log file captured by firewall system

SRC-IP	DST-IP	SRC-PORT	DST-PORT	SIZE
192.168.65.138	64.4.11.36	1520	80	136
192.168.65.138	68.232.44.119	1520	80	153
192.168.65.138	68.232.44.119	1035	53	135
192.168.65.138	58.26.1.16	1537	80	136
192.168.65.138	58.26.1.16	1035	53	126
192.168.65.138	239.255.255.250	68	67	53
192.168.65.138	239.255.255.250	1032	1900	55
192.168.65.138	239.255.255.250	68	67	128
192.168.65.138	192.168.65.253	1079	1900	54

Figure 6: specified entries taken by the proposed system for filteration

SRC-IP	DST-IP	SRC-PORT	DST-PORT	SIZE
192.168.65.138	64.4.11.36	1520	80	136
192.168.65.138	68.232.44.119	1520	80	153
192.168.65.138	68.232.44.119	1035	53	135
192.168.65.138	58.26.1.16	1537	80	136
192.168.65.138	58.26.1.16	1035	53	126
192.168.65.138	239.255.255.250	68	67	53
192.168.65.138	239.255.255.250	1032	1900	55
192.168.65.138	239.255.255.250	68	67	128
192.168.65.138	192.168.65.253	1079	1900	54

Figure 7: highlighted entries are eliminated by the rule base

```
IP :- 127.0.0.1
127.0.0.1
12700000001
4
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 126.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 339.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 345.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 328.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 161.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 161.0
Intrusion detected from IP:239.255.255.250 from port:1900
size * weight --- 161.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 328.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 345.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 333.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 345.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 333.0
Intrusion detected from IP:255.255.255.255 from port:67
size * weight --- 328.0
```

Figure 8: final list of IP addresses generated by GA-RIDS

VI. CONCLUSION AND FUTURE SCOPE

In this paper we have successfully evolved the rule set which can detect existing as well as new intrusions. So as the result generated; the system can be integrate with any of the IDS system to improve the efficiency and the performance. The system can also be able to integrate to the input to the

firewall system. In this paper, we have discussed the GA processes and evolution operators also discussed the overall implementation of GA into proposed system. The various operators like selection, crossover and mutation is also discussed.

In proposed system we are applying single filtration to the system but in future our plan is to apply multiple filters to enhance the system performance and to reduce time complexity of execution. Again we are planning to apply the proposed system output to the security system like Firewall machine to block the traffic whose IP address entries are made.

REFERENCES

[1] T. Lunt, A. Tamaru, F. Gilham, R. Jagannathan, P. Neumann, H. Javitz, A. Valdes, and T. Garvey. "A real-time intrusion detection expert system (IDES)" - final technical report. Technical report, Computer Science Laboratory, SRI International, Menlo Park, California, February 1992.

[2] K. Ilgun, R. A. Kemmerer, and P. A. Porras. "State transition analysis: A rulebased intrusion detection approach". IEEE Transactions on Software Engineering, 21(3):181–199, March 1995

[3] John E. Dickerson, and Julie A. Dickerson "Fuzzy Network Profiling for Intrusion Detection" Electrical and Computer Engineering Department Iowa State University Ames, Iowa, 50011.

[4] Rui Zhong, and Guangxue Yue "DDoS Detection System Based on Data Mining" ISBN 978-952-5726-09-1 (Print) Proceedings of the Second International Symposium on Networking and Network Security (ISNNS '10)Jinggangshan, P. R. China, 2-4,April.2010,pp.062-065.

[5] Dietrich, S., Long, N., and Dittrich, D. 2000. Analyzing distributed Denial of service attack tools: The shaft case. In Proceedings of 14th Systems Administration Conference. New Orleans, Louisiana, USA, 329-339.

[6] Wei Li "Using Genetic Algorithm for network intrusion detection"

[7] B.Upalhaiah, K. Anand, B. Narsimha, S. Swaraj, T. Bharat, "Genetic Algorithm Approach to Intrusion Detection System" ISSN: 0976-8491 (online) | ISSN : 2229-4333 (print), IJCST VOL3, ISSUE 1, JAN-MARCH 2012.

[8] Shrinivasa K G, Saumya chandra, Sidharth Kajaria, Shilpita mukharjee, "IGIDS: Intelligent intrusion detection system using Genetic Algorithm", 978-1-4673-0126-8/11/2011 IEEE.

[9] Anup Goyal, Chetan Kumar, "GA-NIDS : A genetic algorithm based network intrusion detection system",

[10] Atul Kamble, "Incremental Clustering in data mining using genetic algorithm", IJCTE, Vol 2, No. 3, June, 2010

[11] Shaik Akbar, Dr. J. A. chandulal, Dr. K. Nageswara Rao, G. Sudheer Kumar, "troubleshooting technique for intrusion detection sytem using genetic algorithm", IJWBC, vol 1(3), december 2011

[12] Suhail Owais, Vaclav Snasel, Pavel Kromer, Ajith abrahim,"Survey: Using genetic algorithm approach in intrusion detection system techniques", 7th computer information system and industrial management applications,2008 IEEE

[13] P. Jongsubsuk "Intrusion Detection System Using Genetic Algorithm",Science and Information Conference 2014 August 27-29, 2014 | London, UK

[14] "Intrusion Detection System using Fuzzy Genetic Algorithm", "International Conference on Pervasive Computing (ICPC)", -1-4799-6272-3/15/$31.00(c) 2015 IEEE

[15] P.Jongsuebsuk, N. Wattanapongsakorn, C. Charnsripinyo , "Real-Time Intrusion Detection with Fuzzy Genetic Algorithm." ©2013 IEEE.

[16] T.P. Fries, "A fuzzy-Genetic approach to network intrusion detection," GECCO'08: The 10th Annual Conference on Genetic and Evolutionary Computation, 2008, pp. 2141-2146.N.

[17] J. Gomez and E. León, "A fuzzy set/rule distance for evolving fuzzyanomaly detectors," IEEE International Conference on Fuzzy Systems ART. No. 1682017, pp. 2286-2292.

[18] W. Lu and I. Traore, "Detecting New Forms of Network Intrusion Using Genetic Programming". Computational Intelligence, vol. 20, pp. 3, Blackwell Publishing, Malden, pp. 475-494, 2004.

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

[20] R. H. Gong, M. Zulkernine, and P. Abolmaesumi, "A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection," IEEE, University Kingston, Ontario, Canada, 2005.

[21] A. Goyal and C. Kumar, "GA-NIDS: A Genetic Algorithm based Network Intrusion Detection System", 2008. S. Sonawane, S. Pardeshi and G. Prasad, "survey on intrusion detection techniques", World Journal of Science and Technology, 2(3):127-133, 2012.