Intrusion Detection System Using Genetic Algorithm-A Review

*Sharmila Devi, **Ritu Nagpal
*M.Tech (CSE) Student, **Assistant Professor
Deptt. Of CSE, Guru Jambheshwar University of Science and Technology,
Hisar-125001, Haryana
soni.sgn@gmail.com, ritu_nagpal22@yahoo.co.in

Abstract: Today we are suffering from many problems because of intruder interference in our communication with other person/organisation. We need a very safe and secure intrusion detection system. So, intrusion detection has become an important area of research The existing systems are not completely flawless and secure. So, there is the need to improve the existing system. In this paper, firstly we are discussing about the existing network intrusion detection system SNORT and its drawback then discuss about different research areas which were taking place to improve the performance of existing system with the help of genetic algorithm.

Keyword: Intrusion Detection System, Genetic Algorithm, Snort, Network attack, Denial of service.

I. INTRODUCTION

This paper discusses how Genetic Algorithm (GA) can be used to improve performance over existing network intrusion detection system. In existing system SNORT rule cannot be created at run time, rules or expected behaviour stored already in ruleset. If the behaviour of network connection deviates from expected normal behaviour which is stored in ruleset, it will be considered as intrusion. But objective of this paper is to generate rules at run time i.e. add the rule in ruleset with time using genetic algorithm.

II. INTRUSION DETECTION SYSTEM

Intrusion detection systems monitor the network resources and sensing whether a system or network is being used by an authorized person. Two general approach of intrusion detection system are: misuse detection and anomaly detection. Misuse detection IDS implements on the basis of pattern /signature. Misuse detection approach classified into: signature based, rule based and data mining based techniques [1]. Anomaly detection defines the normal behaviour of system. Behaviour change from normal is considered as intrusion [10]. There are two types of intrusion detection system: passive and reactive. In passive system IDS sensor detects the intrusion, log the information and signal the alert on console. In reactive system/intrusion prevention system auto response to the suspicious activity by resetting the connection or by reprogramming the firewall to block network traffic from malicious source [10].Intrusion detection system model is independent of any particular system application environment, system vulnerability or type of intrusion. Six main component of IDS model: subject, object, audit record, profile, anomaly record and activity rule [10]. Intrusion detection system requires the use of metrics. A metric is a random variable x representing a quantitative measure over time. These metrics are of three types: event counter, time interval and resource measurement [10].The use of expert system in IDS was very significant in development of effective detection based information security system. Expert system consists of set of rules which encode the knowledge of human "expert". These rules are used by system to make decision. Expert system permits the incorporation of human experience into computer application which is used to identify activities which matched with the defined characteristic of misuse and attack [5].

Types of intrusion detection system (IDS):

- Network based IDS: This detect the intrusion like denial of service, port scan, and crack into computer by monitoring network traffics.
- Host based IDS: It will search inside the system, not outside. It collects information of individual computer such as web server.
- Protocol based: It is installed on web server and analysis of protocol, monitoring the traffic between a connected device and system it is protected.[21]
- An IDS is composed of several composed of several components:
- Sensor: generate security events.
- Console: monitor event, alerts and control sensors.
- Engine: record events logged by sensors in a data base [6].

III. SNORT

SNORT is an open source ID that is used on Window or Linux operating system. Snort is rule based detection engine which is freely available. Snort is capable of performing real time traffic, analysis, packet logging on IP network. It can be detect variety of attack [6].

By protocol analysis and content searching, snort detects thousand of worms, vulnerability exploit attempts, port scan and other behaviour [8][9]. Snort is configurable in three modes: sniffer mode, packet logger mode, network intrusion detection system mode. In sniffer mode it simply reads the packets of network and display them on screen. To print the TCP/IP packet to screen in this mode just type:

```
. /snort -v
```

If you want to see application data in transit try:

```
. /snort -vd
```

For more detail type:

```
. /snort -vde
```

In Packet logger mode record the packet to the disk for this type:

```
. /snort –dev –l. /log
```

Network intrusion detection system mode analyzes the network traffic against a user –defined rule set:

```
/snort –dev –l ./log –h 172.18.7.217 –c snort.conf
```

Snort .conf is the snort configuration file [7].

Each snort rule has two part header and content of data packet. A snort rule is:

Alert tcp any any -> 172.18.7.218 111

(content:” idc|3a3b|”;msg:”mountd access”;)

Here is some example of snort rules:

Blacklist rule:

alert tcp $HOME_NET any -> $EXTERNAL _NET$HTTP_PORTS(msg: "Blacklist uri request for known malicious URI - /message. php?subid="; flow:to_server, established; content: "/message .php? subid="; nocase; http_uri; content: "version=_nn2"; nocase; http_uri; metadata: impact_flag red; reference:url, labs.snort.org /docs/ 16925.html; classtype:trojan-activity; sid:16925; rev:4;)

DOS Rule:

alert udp any 19 <> any 7 (msg:"DOS UDP echo + chargen bomb"; flow: to_server; reference:cve, 1999-0103; reference:cve,1999-0635; classtype: attempted-dos; sid:271; rev:9;)

Exploit Rule:

alert ip $EXTERNAL_NET any -> $HOME_NET any (msg:"EXPLOIT Linux SCTP malformed forward-tsn chunk arbitrary code execution attempt"; ip_proto: 132; content:"|C0 00|"; depth:2; offset:12; byte_test : 2,>,500,0,relative,big; metadata: policy balanced-ips drop, policy security-ips drop; reference: bugtraq, 33113; reference:cve,2009-0065; classtype:attempted-admin; sid:15490; rev:3;)

Deleted Rule:

alert tcp $EXTERNAL_NET any -> $HOME_NET 443 (msg:"DELETED WEB-MISC TLSv1 Client_Certificate handshake"; flow:established, to_server; ssl_version:tls1.0; content:"|16 03 01|"; content:"|0B|"; within:1; distance:2; flowbits: set,tlsv1.client_hello.certificate; flowbits:noalert; class type:protocol-command-decode; sid:17748; rev:5;)

IV.PROBLEM IN EXISTING SYSTEM

Snort performance evaluation:

F.Alserhana at el. [20] evaluated the performance of snort in high speed network. They were having two cases: 1) Snort and attacker on different operating system platforms. 2) Snort and attacker on same operating system platform. In the first case snort only detect 20% of attack at 1.0 gbps speed of network traffic, when snort installed on Window XP SP2 and was generated from Linux 2.6. In second case snort detect 100% attack up to 400 mbps but only capture 30% of attack at 1.0 gbps. CPU utilization by snort is 80% at 500 mbps input traffic but only 30% at 1.0 gbps. So in this way performance of snort degrade as we increase network traffic. Different problem arise in existing system [14]. These problems are: Fidelity problem is caused when data packets traverse through long path and it can be modified by an attacker. Resource usage problem caused because component of IDS has to be run whole time while there is no intrusion occurred. Reliability problem occurred because the component of intrusion detection system is implemented as separated programs, they are susceptible to tempering and an intruder can disable or modify them. So to overcome these problems we are using Genetic Algorithm [14].

V. GENETIC ALGORITHM

It is not technically feasible to build a system which is having no vulnerabilities. So, intrusion detection has become an important area of research. If an intrusion slightly deviates from the already defined pattern then it will consider as normal and if normal behaviour slightly changes it may be treated as intrusion. Intrusion detection system offer many techniques which recognize and differentiate between normal and intrusion data. Genetic algorithm can be used to tune the membership function of IDS [4].

Genetic Algorithm is a family of computational model based on principles of evolution and natural selection. GA convert the problem into a model by using chromosomes like data structure and evolve the chromosomes using selection, recombination and mutation operator [11][12].

GA begins with randomly selected population of chromosomes which represents the problem to be solved. An evaluation function is used to calculate the “goodness” of each chromosome. The operation start from an initial population of randomly generated chromosomes population evolved for a number of generation and every time quality of an individual gradually improved. Three basic GA operator are applied to each individual i.e. selection, crossover and mutation.

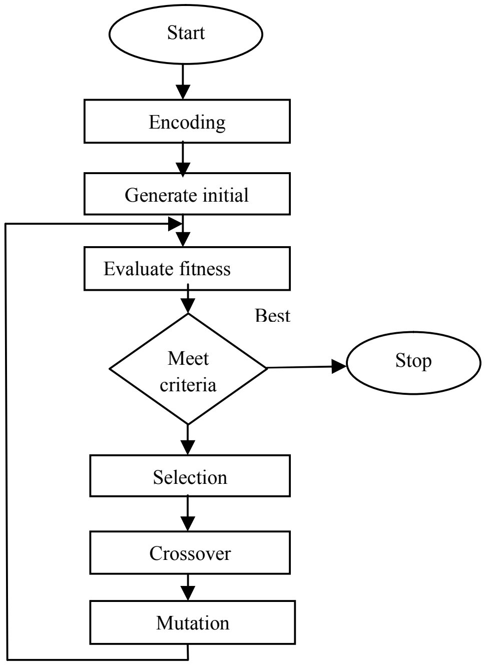

Figure1.Genetic algorithm process [4]

Firstly a number of individual are selected based on user defined fitness function, the remaining are discarded. Next, a number of individual are selected and paired with each other. Each pair produces one offspring by applying crossover operator. At the end a certain number of individual are selected and mutation operator applied i.e. a randomly selected gene of individual abruptly changes its value [13].

VI. RELATED WORK

We have concluded from the previous research [14] that there are three factor of genetic algorithm: 1) Fitness function 2) Representation of individual 3) GA parameters. Genetic Algorithm based intrusion detection system divided into two parts: precalculation phase and intrusion detection phase. In precalculation phase, a set of chromosomes is created using training data in offline environment. [13] In intrusion detection phase, the generated rules are used to classify incoming network connections in real time environment using evaluation process i.e. selection, crossover and mutation [13]. After generating rule it is easy to detect intrusion. Precalculated data is used in this phase to find out fitness of each chromosome. If a better equation is used in these detection process false positive rates will be much slower [14].

Seven network features are selected to form a classification rule. These features are: duration, protocol, source_port, destination_port, source_ip, destination_ip and attack_name [13]. In the real world the types of intrusion change and become complicated very rapidly. So, proposed detection system can upload and update new rules to the system. It is cost effective and adaptive [13].GA can be used to generate the rule for detecting normal and anomalous connections. These rules are stored in ruleset in the form of if {condition} then {act}. Condition part check for matching the current network connection and rules in the ruleset if any connection having same source IP address, destination IP address, destination port number and connection time then this connection will be stop because it matches with the blacklisted IP address. Final goal of applying GA is to generate rule that match only anomalous connection [15].

These rules are tested on historical connection and used to filter the new connection. This paper presents that implementation of GA is unique as it consider both temporal and spatial information of network connection during encoding the problem [15].

If any function produces more and more new rule then add them to the existing rules. Functional advantage is that every time new rules generate, the number of rules overall doubles. So administrator has no need to keep account of all these rules [16].

In genetic algorithm three sub problem arise: 1) coding the chromosomes 2) selecting GA operator 3) find a valid fitness function. Key idea of this approach is automatically construct the rule of ID. This approach implements the heuristic search in the space of network information and finds the same type of attack [17].

Architecture of genetic algorithm for IDS:

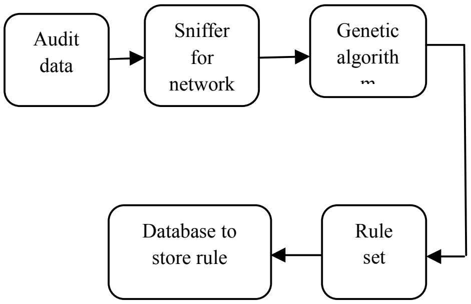

Figure2.Genetic Algorithm for IDS [22]

It requires collecting network data for audit which contains normal and abnormal data. After collecting data , network sniffer will analyze the data and will send it to genetic algorithm. After applying fitness function, rules are added to rule set which are stored in rule base [22].

GP work on the population of parse tree, which is made up of internal nodes and leaf nodes. Internal nodes are called primitive functions and leaf nodes are the terminal [18]. Input of the program is given by terminal, which are the independent variable and set of constant. The rule generated by using GP is in the form "if antecedent then consequent". Three genetic operators are used crossover, mutation and dropping condition. New rule evolved using dropping are like this:

if condition1 and condition2 the consequence can be change to

if condition1 and any then consequence.

A removal approach is also introduced to the simulated artifacts attribute because of over optimistic evaluation of network anomaly detectors. New system outperform over existing system. Limitation of GP is that the algorithm needs two passes during training, resulting in the inefficiency of detector [18].

[19] Binary code has a continuous function of discrete mapping error for some multidimensional and high precision requirement of continuous function optimization. They use gray code coding because it improve the GA local search capability. Selection of population size is one of the most important parameter in GA. If the population size is too large then it will lower the efficiency of GA. If it is too small, it may improve the speed but lower the diversity of population and it will cause premature conversion [19]. This system created lot of selection methods such as random search, adjacent search, multipoint search and best individual multipoint search because the existing selection operator roulette wheel leads to "premature conversion" and slowdown search process. They only use three point crossovers; position can be randomly selected with no repetition. By using improved GA their work effectively improved the ID rate [19].

VII. CONCLUSION

In described techniques, Genetic Algorithm decreases the false +ve rate. Proposed detection system uploads and update new rule to the system. Implementation of Genetic Algorithm is unique as it considers both temporal and spatial information during encoding the problem. New rules are generated at run time, so administrator has no need to keep track of all these rules.

REFERENCE

[1]. Aleksandar L., Vipin K., and Jaideep S., Massive Computing Managing Cyber Threats, Issues, Approaches, and Challenges. Intrusion Detection: A Survey: Computers/General Information. Springer, 2005.

[2]. Denning, Dorothy (1986) an intrusion detection model .IEEE Transaction on software Engineering, vol.se-13, no.2.

[3]. McHugh, John, 2001. "Intrusion and Intrusion Detection." Technical Report. CERT Coordination Center, Software Engineering Institute, Carnegie Mellon University.

[4]. S.Owais, V.Snasel, A.Abraham,"Survey: Using Genetic Algorithm Approach in Intrusion Detection Systems Techniques" 7th Computer Information Systems and Industrial Management Applications IEEE, 2008.

[5]. Anderson, D., Frivold, T. & Valdes, A. (May, 1995). Next -generation Intrusion Detection Expert System (NIDES): A Summary. SRI International Technical Report SRI-CSL-95-07.

[6] K. Scarfone, and P. Mell, ―Guide to Intusion Detection and Prevention Systems‖, National Institute of Standards and Technology NIST. Computer Security, 2007.

[7]. Snort Users Manual, http://www.snort.org.

[8]. Sectools.Org:2006Results;http://sectools.org/tools2006.html

[9].SecTools.Org: Top 125 Network Security Tools; http://sectools.org/tag/ids/.

[10]. en.wikipedia.org/wiki/intrusion_detection system.

[11]. Sinclair, Chris, Lyn Pierce, and Sara Matzner. 1999. "An Application of Machine Learning to Network Intrusion Detection." In Proceedings of 1999 Annual Computer Security Applications Conf. (ACSAC), pp. 371-377. Phoenix, Arizona. URL: http://www.acsac.org/1999/papers/fri-b-1030-sinclair.pdf (30 Oct. 2003).

[12]. Whitley, Darrell. 1994. "A Genetic Algorithm Tutorial." Statistics and Computing 4: 65-85.

[13]. R. H. Gong, M. Zulkernine, and Purang, ―A software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection‖, SNPD/SAWN'05, IEEE, 2005.

[14]. M.S. Hoque, M.A. Mukit, M.A.N. Bikas "An implementation of intrusion detection system using Genetic Algorithm" International Journal of Network Security & its Application (IJNSA),Vol.4,no.2,march 2012.

[15] W. Li, "Using Genetic Algorithm for Network Intrusion Detection", Proceedings of the United States Department of Energy Cyber Security Group, 2004.

[16]. Ch. S.Keerthi.N.V.L, P.L.prasanna, B.M. Priscilla, M.V.B.T.santhi "Intrusion Detection System Using Genetic Algorithm" International Journal of P2P Network Trends and Technology-Volume1 issue2-2011.

[17]. Guan Jian,Liu Da-Xin, Cui Bin-Ge " An Induction Learning Approach for Building Intrusion Detection Models Using Genetic Algorithem" proceeding of IEEE 5th world conf. On Intelligent Control and Automation.June 15-19.2004,P.R.China.

[18]. Chuanhuan Yin, Shengfeng Tian, Houkuan Huang and Jun He " Applying Genetic Programming to Evolve Learned Rules for Network Anomaly Detection" L. Wang, K. Chen, and Y.S. Ong (Eds.): ICNC 2005, LNCS 3612, pp. 323 – 331, © Springer-Verlag Berlin Heidelberg 2005.

[19] QIAO Pei-li, CHEN Shi-feng,SU jie " The Research of NIDS Based on Improved GA" Proceeding of IEEE conference,2009.

[20]. F. Alserhani, Monis Akhlaq, I. U. Awan, A. J. Cullen, J. Mellor ,Pravin Mirchandani "Snort Performance Evaluation" Informatics Research Institute, University of Bradford, Bradford, BD7 1DP, United Kingdom.

[21]. J. P. Planquart, "Application of Neural Networks to Intrusion Detection", SANS Institute Reading Room.

[22]. Vivek K. Kshirsagar, Sonali M. Tidke & Swati Vishnu "Intrusion Detection System using Genetic Algorithm and Data Mining: An Overview" International Journal of Computer Science and Informatics ISSN (PRINT): 2231 –5292, Vol-1, Iss-4, 2012