# A SURVEY ON GENETIC ALGORITHM FOR INTRUSION DETECTION SYSTEM

MS. DIMPI K PATEL
*Department of Computer Science and Engineering,*
*Hasmukh Goswami college of Engineering,*
*Ahmedabad, Gujarat*

**ABSTRACT**

*The Internet has become a part of daily life and an essential tool today. Internet has been used as an important component of business models. Therefore, It is very important to maintain a high level security to ensure safe and trusted communication of information between various organizations.*

*Intrusion Detection Systems have become a needful component in terms of computer and network security. Intrusion detection is one of the important security constraints for maintaining the integrity of information. Intrusion detection systems are the tools used for prevention and detection of threats to computer systems. Various approaches have been applied in past that are less effective to curb the menace of intrusion.*

*In this paper, a survey on applications of genetic algorithms in intrusion detection systems is carried out.*

**Keyword:** *Intrusion Detection System (IDS), Network Intrusion detection system (NIDS), Host Intrusion Detection Systems (HIDS), Genetic algorithm (GA).*

---

## 1. INRODUCTION

The computer technology has evolved at a very fast pace since last few decades. With this rapid growth of computer technology has resulted in the transfer of more and more services to computer based systems. The dependency of such services on computer technology has resulted in the increase of computer related threats. Viruses, Worms and Trojan horses and hacking are some of the major attacks becoming painful for any network systems.

Most of the security mechanisms designed so far, try to prevent unauthorized access to system resources and data. However, designed such systems are not able to completely prevent intrusions into computer systems. The main need is to detect intrusions efficiently and from that their impact can be realized and damages can be repaired. Thus, Intrusion Detection System (IDS) has become one of the hottest research areas in Computer Security now a days.

The conventionally used security tools like firewalls, intrusion detection systems (IDS) has become with supreme significance. Intrusion Detections Systems (IDS) is a new path of security systems, which provides efficient approaches to secure computer networks. Artificial Intelligence approaches have been used at large level to produce a lot of IDS. Some of these approaches dependant on Genetic Algorithms to provide the network with an efficient classifier to recognize and detect intrusions actions.

## 2. SYSTEM OVERVIEW

### 2.1 Intrusion Detection Systems

An intrusion detection system is a device or software application which monitor network or system activities for malicious activities and providing reports to a management stations[1]. The main aim of Intrusion Detection Systems (IDS) is to protect the availability, confidentiality and integrity of critical networked information systems.

Intrusions Detection can be classified into two major categories as network intrusion detection systems (NIDS), host intrusion detection systems (HIDS). In Network Intrusion Detection System, it evaluate information captured from network communications and analyze the stream of packets which travel across the network. In Host Intrusion Detection System, It evaluate information found on a single or multiple host systems, including contents of operating systems, system and application files.

An intrusion detection system generally composed of three functional components [6].

The first component is also known as the 'event generator', is a data source. Data sources can be categorized into four categories namely Host-based monitors, Network-based monitors, Application-based monitors and Target-based monitors.

The second component of an intrusion detection system is also known as the 'analysis engine'. This component takes information from the data source and it examines the data for symptoms of attacks or other policy violations. The analysis engine can use one or both of the following analysis approaches:

I. Misuse/Signature-Based Detection: This type of detection engine detects intrusions which follows well-known patterns of attacks (or signatures) that exploit known software vulnerabilities [9]. The main demerit of this approach is that it only looks for the known weaknesses and may not care about detecting unknown future intrusions .

II. Anomaly/Statistical Detection: An anomaly based detection engine will search for something rare or unusual [13]. They analyses system event streams, using statistical techniques to find patterns of activity that appear to be abnormal. The main demerits of this system are that they are highly expensive and they can recognize an intrusive behavior as normal behavior because of insufficient data.

The third component of an intrusion detection system is the response manager. In basic terms, the response manager will only act when inaccuracies (possible intrusion attacks) are found on the system, by informing someone or something in the form of a response.

Figure 1[14], gives a generic architecture of an intrusion detection system.

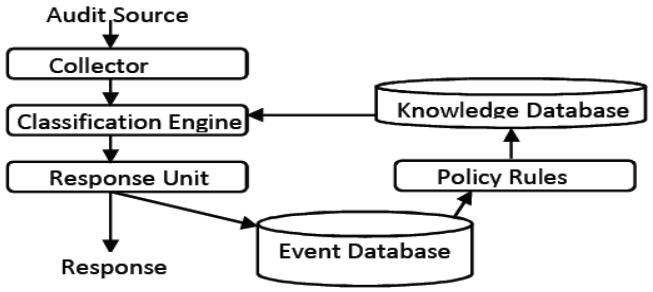


***Fig. 1. Generic Architecture of Intrusion Detection System***

In figure 1, the audit source represents the input to the intrusion detection system. The format of input data can be of different types depending upon the type and location of the intrusion detection system. The collector samples and preprocesses the audit source data. The data is transformed into a standard format known to the internal components of the intrusion detection system. The knowledge database contains information about attacks. The classification engine determines the legitimateness of the received data by comparing it with the attack information stored in the knowledge database. The policy rules are used to configure the response and detection of intrusion system. The response unit produces different types of responses depending upon the incoming events and their severity. The event database stores the detailed information about the events, which is used for various purposes like attack report generation, and framing new rules.

**2.2 Fuzzy logic**

Fuzzy Logic was introduced as a means to the model of uncertainty of natural language. And due to the uncertainty nature of intrusions fuzzy sets are strongly used in discovering attack events and reducing the rate of false alarms at the same time. Basically, intrusion detection systems distinguish between two distinct types of behaviors, normal and abnormal, which create two distinct sets of rules and information. Fuzzy logic could create sets that have in-between values where the differences between the two sets are not well defined. In this case the logic depends on linguistics by taking the minimum of set of events or maximum instead of stating OR, AND or NOT operation in the if-then-else condition. This feature strongly participates in reducing the false positive alarm rates in the system. Applying fuzzy methods for the development of IDS yield some advantages, compared to the classical approach. Therefore, Fuzzy logic techniques have been employed in the computer security field since the early 90's. The fuzzy logic provides some flexibility to the uncertain problem of intrusion detection and allows much greater complexity for IDS[3].

**2.3 Genetic Algorithms**

"A Genetic Algorithm (GA) is a programming technique that mimics biological evolution as a problem-solving strategy"[12]. It is based on Darwinian's principle of evolution and survival of fittest to optimize a population of candidate solutions towards a predefined fitness[2].

The process of a genetic algorithm usually begins with a randomly selected population of chromosomes. These chromosomes are representations of the rules. According to the attributes of the rules, different positions of each chromosome are encoded in binary bits. These positions are sometimes referred to as genes and are changed randomly within a range during evolution. The set of chromosomes during a stage of evolution are called a population. An evaluation function is used to calculate the "Fitness" of each chromosome. During evaluation, two basic operators, crossover and mutation, are used to simulate the natural reproduction and mutation of species. The selection of chromosomes for survival and combination is biased towards the fittest chromosomes.

The following figure 2 taken from [2] shows the structure of a simple genetic algorithm. Starting by a random generation of initial population, then evaluate and evolve through selection, crossover, and mutation. Finally, the best individual (chromosome) is picked out as the final result once the optimization meet its target.

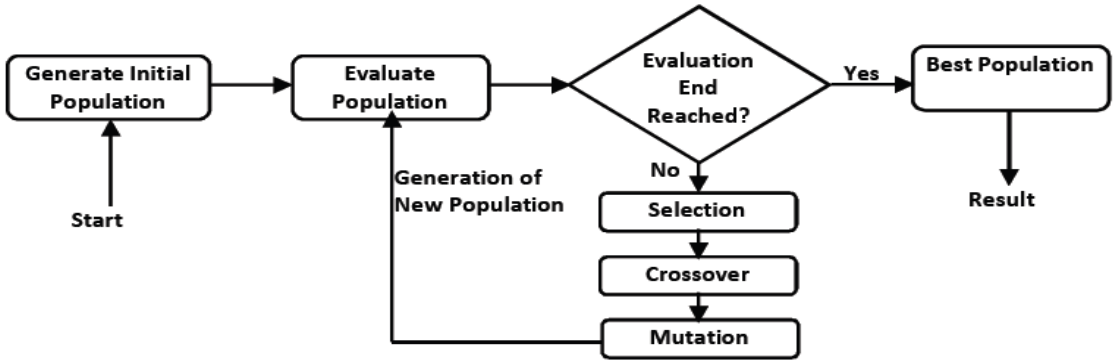


***Fig. 2. Structure & Processing in a Genetic Algorithm.***

**Major Steps in Genetic Algorithm**

**Algorithm:** Rule set generation using genetic algorithm.

**Input:** Network audit data, number of generations, and population size.

**Output:** A set of classification rules

1. Initialize the population

2. Check the fitness function

3. Select only those rules that that meets the fitness criteria.

4. Perform crossover for reproduction of new rule by exchanging some bits

5. Perform mutation by flipping some bits

6. Again go to line 2, until the specified numbers of rules are not generated.

## 3. LITERATURE REVIEW

Several case studies and experiments were applied regarding the use of genetic algorithm in intrusion detection system. Some import applications of soft computing techniques for network intrusion detection is described in this section. Several Genetic Algorithms and Genetic Programming has been used for detecting intrusion detection of different kinds in different scenarios.

In [5], Chittur A applied detection of computer intrusions and malicious computer behavior and analyzed the effectiveness of a Genetic Algorithm.

In [2], Li described a method using GA to detect anomalous network intrusion . His approach includes both Quantitative and categorical features of network data for deriving classification rules. He addressed the Factors affecting to Genetic Algorithm are in detail. The inclusion of quantitative feature can increase detection rate but experimental results are not available.

In [4], Xia et al. detected anomalous network behaviors based on information theory by using GA. Some network features can be identified with network attacks based on mutual information between network features and type of intrusions and then using these features a linear structure rule and also a GA is derived. The advantage of the approach of using mutual information and resulting linear rule seems very effective because of the reduced complexity and higher detection rate. The approach has disadvantage that it considered only the discrete features.

In [11], Abdullah showed A GA based performance evaluation algorithm to network intrusion detection. The traffic data was filtered by information theory in his approach.

In [10], Lu and Traore used support-confidence framework as the fitness function and accurately classified several network intrusions. Disadvantage of their approach is that, use of genetic programming made the implementation procedure very difficult and also for training procedure more data and time is required.

In [8], Crosbie and Spafford applied the multiple agent technology and GP to detect network anomalies. For both agents they used GP to determine anomalous network behaviors and each agent can monitor one parameter of the network audit data. Advantage of this technique is that it acts well when many small autonomous agents are used. And Disadvantage is if the agents are not properly initialized the training process can be time consuming, when communicating among the agents.

In [7], Gong presented An implementation of GA based approach to Network Intrusion Detection using GA and showed software implementation. In this approach he derived a set of classification rules using a support-confidence framework to judge fitness function.

## 4.CONCLUSION

In this survey an introduction to intrusion detection and genetic algorithms are presented. It mainly focuses on various IDS models. It is realized that various techniques can be used to implement IDS. The paper provides enough information for newcomers to the field of genetic algorithms based intrusion detection.

**5.ACKNOWLEDGEMENT**

I want to thank my supervisor Prof. Indr Jeet Rajput, Assistant Professor in HGCE, Ahmadabad not only for his continued support but for the motivation and fruitful advises in accomplishing this task.

## 6.REFERENCES

[1] http://en.wikipedia.org/wiki/Intrusion_detection_system

[2] W. Li, “Using Genetic Algorithm for Network Intrusion Detection”. SANS Institute, USA, 2004.

[3] N. Bashah Idris, B. Shanmugam “Novel Attack Detection Using Fuzzy Logic and Data Mining”, International Conference of Soft Computing and Pattern Recognition, SOCPAR '2009.

[4] T. Xia, G. Qu, S. Hariri, M. Yousif, "An Efficient Network Intrusion Detection Method Based on Information Theory and Genetic Algorithm", Proceedings of the 24th IEEE International Performance Computing and Communications Conference (IPCCC '05), Phoenix, AZ, USA, 2005.

[5] A. Chittur, "Model Generation for an Intrusion Detection System Using Genetic Algorithms". January 2005.

[6] R. G. Bace, "Intrusion Detection", Macmillan Technical Publishing, 2000.

[7] R. H. Gong, M. Zulkernine, P. Abolmaesumi, "A Software Implementation of a Genetic Algorithm Based Approach to Network Intrusion Detection", 2005.

[8] M. Crosbie, E. Spafford, "Applying Genetic Programming to Intrusion Detection", Proceedings of the AAAI Fall Symposium, 1995.

[9] S. Kumar, E. Spafford, "A Software architecture to Support Misuse Intrusion Detection", in the 18th National Information Security Conference, pp. 194-204, 1995.

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

[11] B. Abdullah, I. Abd-alghafar, Gouda I. Salama, A. Abd-alhafez, "Performance Evaluation of a Genetic Algorithm Based Approach to Network Intrusion Detection System", 2009.

[12] Md. Sazzadul Hoque, Md. Abdul Mukit , Md. Abu Naser Bikas, "An Implementation of Intrusion Detection System using Genetic Algorithm", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012

[13] S. Kumar, "Classification and Detection of Computer Intrusions", Purdue University, 1995.

[14] M. Arvidson and M. Carlbark, "Intrusion Detection Systems: Technologies, Weaknesses, and Trends," 2003.