Multi Level Malicious Behaviour Detection and Analysis Using Genetic Algorithm for Social Network

Sudhansu Gaur And Ashwani Kumar

Computer Science Engineering Department Guru Jambheshwar University Science &Technology, Hisar (Haryana) , India

Abstract

Protect sensitive information of different social network. We propose taking benefits of real-life social trust between different users as well as threshold cryptography. This leads to a new type of complex systems, the increase in network based attacks in general, and the world-wide access to computer networks and systems in particular, those responsible for network and computer system security need to utilize every tool available.

First, in this paper will examine detailed implementation of applying genetic algorithm to intrusion detection Genetic Algorithm(Fitness ,Fitness_threshhold,X,Y,Z) Then, genetic algorithms will be discussed. After this, the Combining of genetic algorithms with intrusion detection will be reviewed. Finally, future steps will be discussed in the use of a genetic algorithm within intrusion detection.

Keywords:- Intrusion Detection; Network Intrusion Detection; Genetic Algorithm;

I Introduction

In recent years, Intrusion Detection System (IDS) has become one of the hottest research areas in Computer Security. It is an important detection technology and is used as a countermeasure to preserve data integrity and system availability during an intrusion. When an intruder attempts to break into an information system or performs an action not legally allowed, we refer to this activity as an *intrusion* (Graham, 2002; see also Jones and Sielken, 2000). Intruders can be divided into two groups, *external* and *internal*. The former refers to those who do not have authorized access to the system and who attack by using various penetration techniques. The latter refers to those with access permission who wish to perform unauthorized activities. Intrusion techniques may include exploiting software bugs and system misconfigurations, password cracking, sniffing unsecured traffic, or exploiting the design flaw of specific protocols (Graham, 2002). An Intrusion Detection System is a system for detecting intrusions and reporting them accurately to the proper authority. Intrusion Detection Systems are usually specific to the operating system that they operate in and are an important tool in the overall implementation an organization's information security policy (Jones and Sielken, 2000), which reflects an organization's statement by defining the rules and practices to provide security, handle intrusions, and recover from damage caused by security breaches.

Intrusion is a term often used to define a set of actions which attempt to compromise the confidentiality, integrity, or availability (CIA) of a target system. Thus, studies on intrusion detection systems (IDSs) are generally aimed at detecting an intruder who tries to break into a computer system or detecting a legitimate user who attempts to misuse the system's resources in real-time. Based on an intrusion detection model, an intrusion detection technique is categorized either as misuse detection or anomaly detection [11]. In misuse detection schemes, the intrusions are defined as attacks on known vulnerability of the computer system or network equipment. Misuse ehaviour can be detected by monitoring their performed actions or matching predefined patterns. Misuse detection can handle or find out the known attacks but cannot discover novel attacks because misuse detection depends on the intrusion's database for known attack patterns or well-defined signatures. On the other hand, in anomaly detection schemes, intrusions are discovered by observing deviations from the monitored behaviours or the system usage logs. A system model or profile is built which contains metrics that are derived from the significant ehaviour in system or the process operation.

Metrics are calculated from available parameters such as average CPU load, number of network connections, number of processes per user, the call status, to name a few. But there are some problems in anomaly detection schemes. For example, anomaly detection techniques use significant difference from user profiles or normal ehaviour profiles. Intruders or hackers can attempt to adjust the system's normal profiles by providing fake data over a period of time. According to the data sources, intrusion detection techniques can further be classified as host-based detection and network-based detection [12]. The host-based intrusion detection system (HIDS) collects data such as CPU load, command log, system calls, etc. Because host-based intrusion detection has no information of network events in lower layers, it usually does not detect network-based attacks. Different operating systems or platforms of hosts require different intrusion detection systems;

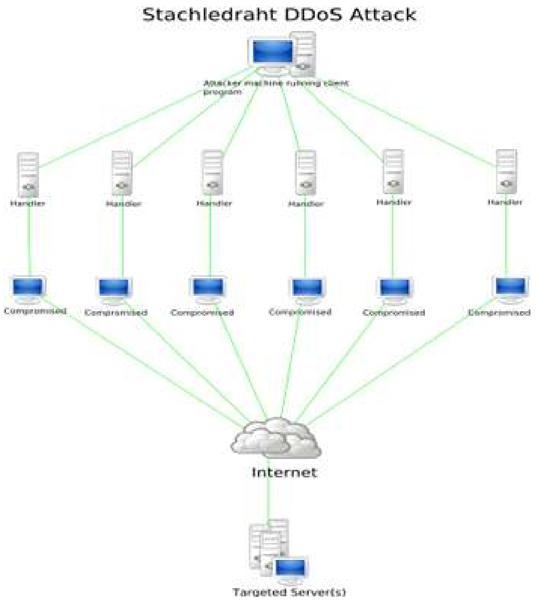


Figure: Denial of Service attack

One such intrusion is a Denial of Service (DoS) attack. A DoS attack is typically an attempt by an attacker to prevent valid users of a service from using or having access to that service (CERT).The goal of intrusion detection is to recognize attempts to sabotage in-place security controls (Berge). Specifically, network traffic is analyzed in search for system based breaches. Network breaches can take various forms.

The paper is organized as follows. First, we start with a description of Security Analysis and Privacy in Section II. In Section III, we describe brief introduction to genetic algorithm. In section IV Security and privacy solution. In Section V, we show how the detailed implementation of applying genetic algorithm to intrusion detection. In Section VI Simulation and Result Finally, in Section VII, presents the conclusion and future work.

II. Security Analysis and Privacy

II a. Security Analysis

Now in this section the writer analyze the security of proposed protocols against threats which could occur during eavesdropping or attacks performed by intruders or attackers. According to [1] in order to meet authentication.

(Prevention or Secure- Not satisfying or partially support) of the first three rows of vulnerabilities and threats assumed for authentication protocols would be extracted and discussed according to investigations applied.

A. Information Leakage

The first vulnerability deliberated in the analysis and in the table is Information leakage which for some protocols or schemes like RHLK, HIDV, SRAC, HBIV, Li et al. and Hung-Yu Chien and Chen Wei Huang there were no implication in references investigated even though stated that most protocols are designed to protect against information leakage and spoofing attack.

B. Spoofing Attack

RHLK is vulnerable to a spoofing attack by impersonating a tag to a legitimate reader. In order to prevent the adversary from performing a Spoofing attack, it is needed that the tag response be randomized in every session, which is not provided in HIDV when it uses a fixed hash value in every authentication Since in LCAP protocol a social network response is randomized in every session, prevention against Spoofing attack.

II b. Why Need Privacy

Privacy is a particularly big concern when billions and billions of small devices are expected to be embedded into goods and to send various information over the air about them and their holders. To achieve this goal, we first have to define precisely the different uses of these devices and the different needs they generate in terms of privacy.

1. Detection Needs

Detection consists, by using a genetic algorithm reader, in first finding objects that emit signals with sufficient power to reach the reader, wherever they may be hidden, and second getting information given by the social network is appropriate, every accounting application can be fulfilled by this procedure.

However, when the level of information which is publicly available from the social network is too high, privacy concerns arise, as provided data could allow anyone to uniquely (oralmost uniquely) identify each host. Therefore, it becomes necessary to design a general scheme for system and readers, which allows tags to disclose the nature of the items.

2. Authentication Needs

Another emerging application of system is control of authenticity. Manufacturers of the luxury industry have already begun to integrate system in their products, so that counterfeiting can be detected more easily.

3. Identification Need

Identification needs are closely related to traceability, which consumers often consider as a threat to their privacy.

However, traceability is required by many applications (shipments, after-sales follow-up…). Thus, in order to protect consumers' privacy, a first step is to prevent social network readers with no special privilege from tracing items. And apply the intrusion detection Genetic Algorithm (GA) is an efficient investigating method used in computing to locate precise or estimated solutions to optimization and search problems [3].

III Introduction to Genetic Algorithm

Genetic Algorithms are categorized as global search heuristics. In GA heuristic search is based on concept of biological evolution In Genetic algorithms iterative mathematical modeling technique is used to find the optimal combinatorial state given a set of parameters of interest. The population evolution process is simulated through genetic programming [4]. A inhabitants of fixed-length is evolved with a GA by employing crossover and mutation operators along with a fitness function that concludes how likely individuals are to reproduce [5]. A testing solution is developed each of which is estimated (to yield fitness) and a new generation is created from the better of them. The procedure is sustained through a numerous generations with the endeavor that the population should evolve to contain a solution which is acceptable.
Following steps are involved in GA application:

1. At the start of a run of a GA, a huge population of random genes is created. Each represents a different solution to the problem at hand. Let's say in the initial population there are N chromosomes then the following steps are repeated until a solution is found.
2. Every chromosome is tested to see how better it is at solving the problem at hand and allocate a fitness score accordingly.
3. From the current population two members are selected. The chance of being selected is proportional to the chromosomes fitness. Normally Roulette wheel selection method is used.
4. Depending on the crossover rate the bits from each selected chromosome are crossover randomly at chosen point.
5. Step through the chosen chromosomes bits and flip dependent on the mutation rate.

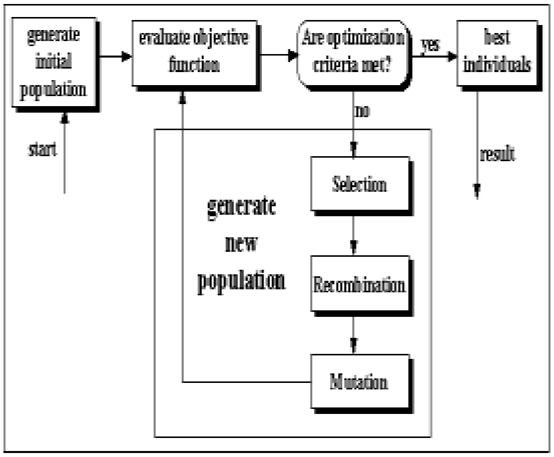


Figure . Structure of a simple genetic algorithm (Pohlheim, 2001)

IV. Security and Privacy Solutions

Security in Wireless Network, we mainly focus on the problem of achieving some of all of the following security contributes or services [6]:

• **Confidentiality**: Confidentiality or secrecy has to do with making information inaccessible to unauthorized users.\
• **Availability**: Availability ensures the survivability of network services to authorized parties when needed despite denial-of-service attacks.
• **Integrity: Integrity** measures ensure that the received data is not altered in transit by an adversary.
• **Authentication**: Authentication enables a node to ensure the identity of the peer node with which it is communicating.
• **Non-repudiation:** Non-repudiation denotes that a node cannot deny sending a message it has previously sent.
• **Authorization:** Authorization ensures that only authorized nodes can be accessed to network services or resources.
• **Freshness**: This could mean data freshness and key freshness. Since all sensor networks provide some forms of time varying measurements, we must ensure each message is fresh. Data freshness implies that each data is recent, and it ensures that no adversary replayed old messages.

V. Proposed Genetic algorithm

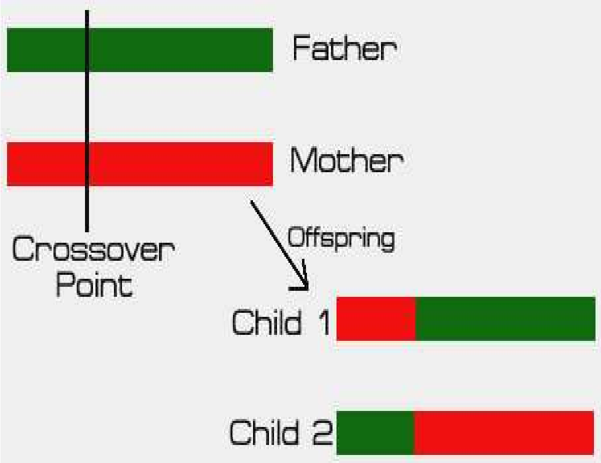
Security and privacy solution in detailed implementation of applying intrusion detection Genetic Algorithm(Fitness ,Fitness_threshhold,X,Y,Z).

```
Step1:-Initialze the value
            X← x  random hypothesis.
Step2:-Evalute the value
             For each f in X, compute fitness(f).
Step3:-Check the value
           While(Maximum Fitness(f)) <
Fitness_Threshold.
Step4:-Select the value
               Probability select (1-Y)X .

Step5:-Crossover:- probability select (Y-X)/2
pairs of hypothesis from x .for each pair
                      (f1,f2), produce two offspring
by applying the crossover operator ,add
                       all  offspring to Xs.
Step6:-Mutate: Invert a randomly selected bit
Z.X random member of Xs.
Step7:-Update : the probability
                        X → Xs.
Step8:-Evalute: for each f in X, compute
fitness(f).
Step9:-Return the hypothesis from X that has the
highest fitness.
Step10:-End
```

Security has to do with making information inaccessible to unauthorized users. Using the operations for genetic algorithm and rule set.

V.a Operations for Genetic Algorithm



For Example

11111000000	Father
00000111111	Mother

Child1-→ 00000000000

Child2-→ 11111111111

V.b The Rule Set

The rule set is produced from the output of the GA
For example, the input of Source IP = 2456233458 (which is an IPv4 address of 209.16.1.10)| Destination IP =2456233456 (which is an IPv4 address of 209.16.1.2) | Destination Port = 2321 | Protocol = 5 | Originator Bytes = 205 | Responder Bytes = 4500 could produce the following rule:

Than Categories as "Attacks"

Else Categories as "General"

VI. SIMULATION AND RESULTS

We simulated intrusion detection genetic algorithm with Simulator NS2.Based on the findings of applying the Network simulator Framework to a theoretical network intrusion, some future steps could be:

- Utilize actual anomaly based network device data to form the input "chromosomes", the gene range values, and the parameters for the evaluation function
- Compare GA results with existing intrusion rule sets for effectiveness.
- Data mine the GA results for patterns or data clusters and then analyze for discoveries.
- Utilize Genetic Programming, which enhance GAs since they produce dynamic programs instead of static chromosomes, which result in more multifaceted and flexible outcomes.

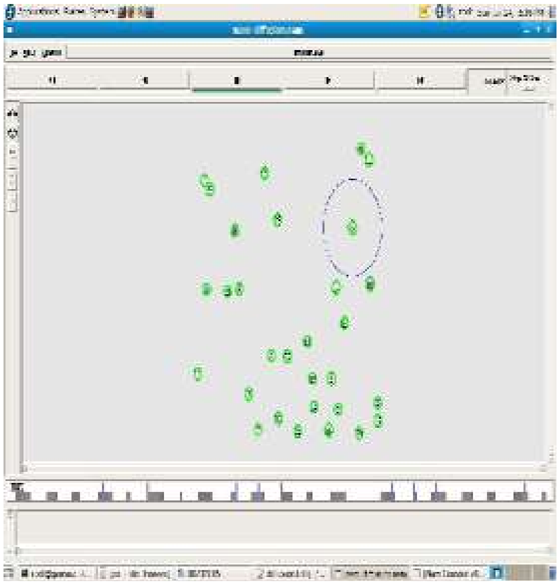


Figure1(a):-Sending the packet

Figure 1(a) is represent the packet, and sender side are all packet using the trust based. And all packet send the receiver side.

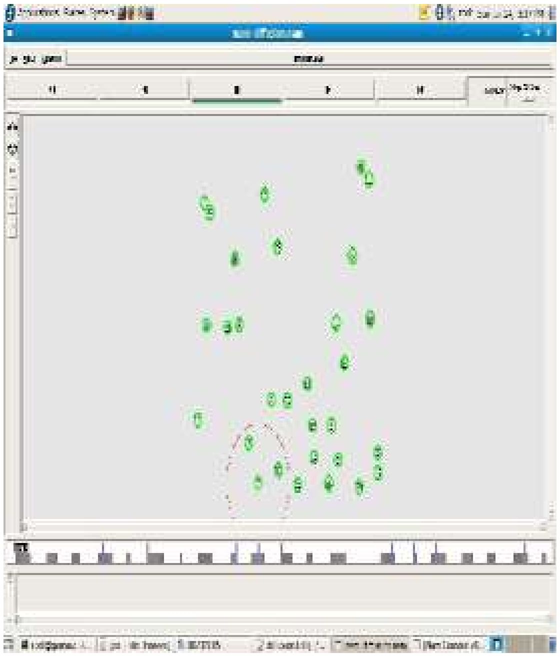


Figure1(b):-Receiving the packet

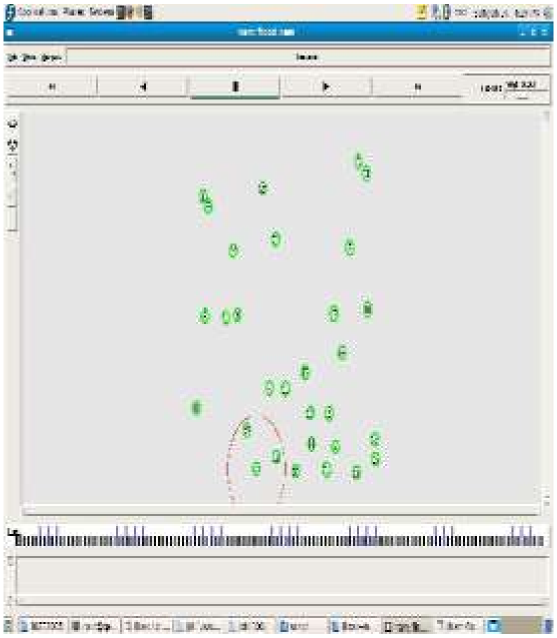


Figure2(a):- Detection for effective packet

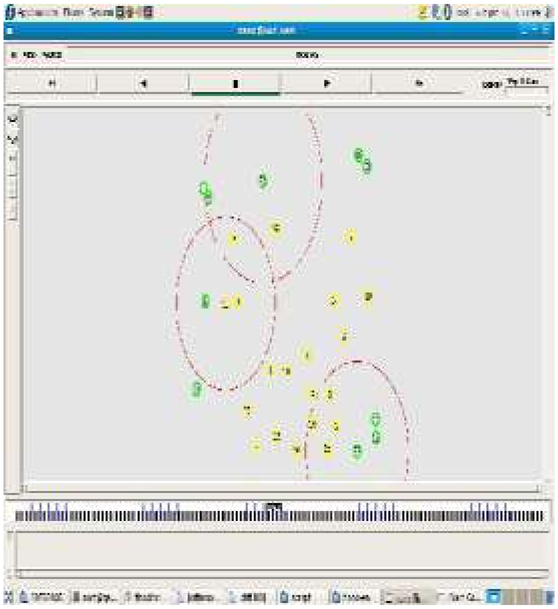


Figure2(b):- The nodes, receive more broadcast became yellow but some information effective packet represent by red circle

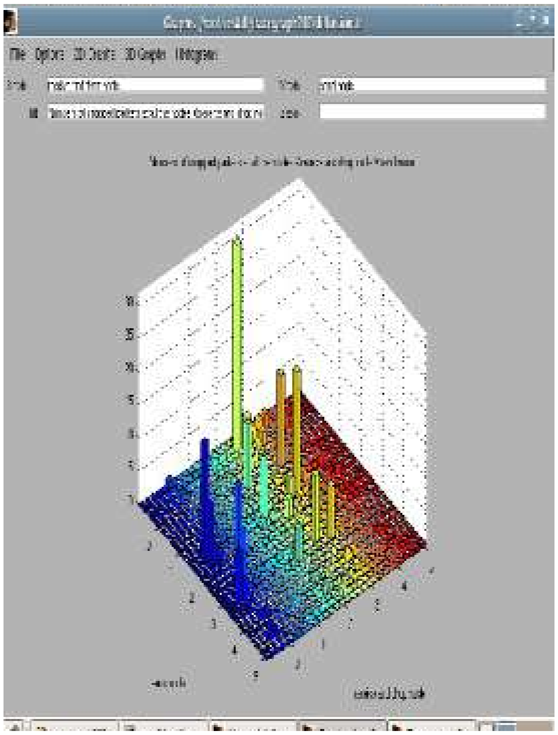


Figure3:- Dropped packed

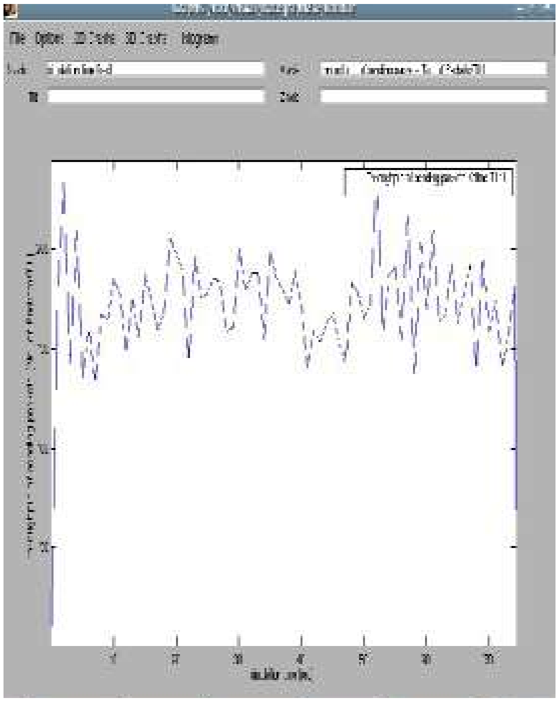


Figure4(a):- Throughput of Sending Packets vs Simulation Time

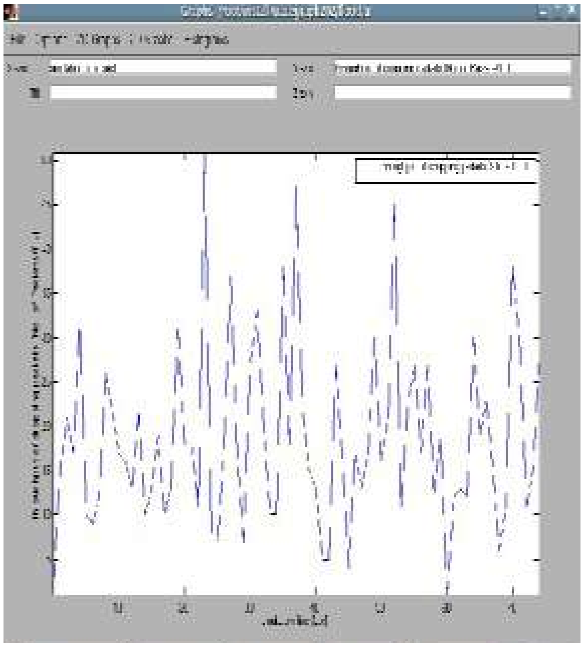


Figure4(b):-Throughput of Receiving Packets vs Simulation Time

Overall output

Serial Number	Occurrence	Performance	Performance
		Training set	Test set
1	100	90%	88%
2	150	84%	81%
3	200	88%	85%
4	125	78%	74%
5	175	77%	75%
6	120	91%	89%

Figure5: Performance of Training Set and Test Set

Also using rule based identification packet and send the IP address then identify true and false node.

IP Address	Identify Node TRUE	Identify Node FALSE	Dropped Packed	Throughput
192.68.1.55	15	5	5	70.5
192.68.1.57	18	2	2	90
192.68.1.58	14	6	6	70
192.68.1.60	17	3	3	80.5
192.68.1.65	16	4	4	80
192.68.1.67	18	2	2	90
192.68.1.70	17	3	3	80.5

Figure6:- Throughput of Receiving Packet

Result

Fixed Packed = 20

Each IP address Sending 20 Packed.

And using Intrusion detection genetic algorithm.

Then identification

Find the True identification packet means as a general packet or original packet or authorized packet

Else

Find the false identification packet means as a attack packet or not original packet or unauthorized packet.

Or

Dropped packed means unauthorized packet.

VI Conclusion and future work

The Genetic Algorithm can be effectively used for formulation of decision rules in intrusion detection through the attacks which are more common can be detected more accurately. This implementation of genetic algorithm is unique as it considers both temporal and spatial information of Rule based classification of DoS and Probe attacks can be used for effective monitoring of the network. Validated through test data network connections during the encoding of the problem; therefore, it should be more helpful for identification of network malicious behaviors.. Beside this, development of knowledge base as a result of GA application can be utilized for further investigation for identification of attribute which contribute for accurate classification of attack.

Future work includes creating a standard test data set for the genetic algorithm proposed in this paper and applying it to a test environment. Detailed specification of parameters to consider for genetic algorithm should be determined during the experiments. Combining knowledge from different security sensors into a standard rule base is another promising area in this work.

REFERENCES

[1] Bezroukov, Nikolai. 19 July 2003. "Intrusion Detection (general issues)." Softpanorama: Open Source Software Educational Society. Nikolai Bezroukov. URL: http://www.softpanorama.org/Security/intrusion_detection.shtml (30 Oct. 2003).

[2] Bridges, Susan, and Rayford B. Vaughn. 2000. "Intrusion Detection Via Fuzzy Data Mining Mining." *In Proceedings of 12th Annual Canadian Information Technology Security Symposium*, pp. 109-122. Ottawa, Canada.

[3] Crosbie, Mark, and Gene Spafford. 1995. "Applying Genetic Programming to Intrusion Detection." *In Proceedings of 1995 AAAI Fall Symposium on Genetic Programming*, pp. 1-8. Cambridge, Massachusetts. URL:

http://citeseer.nj.nec.com/crosbie95applying.html (30 Oct. 2003).

[4] Graham, Robert. Mar. 21, 2000. “FAQ: Network Intrusion Detection Systems.” RobertGraham.com Homepage. Robert Graham. URL: http://www.robertgraham.com/pubs/network-intrusion-detection.html (30 Oct. 2003).

[5].Fiskiran, A.M; Lee, R.B., “Runtime Execution Monitoring to detect and prevent malicious code execution”, Computer Design: VLSI in computers and Processors 2004, IEEE International Conference Pages 452-457.

[6].Baoyi Wang; Feng Li; Shaomin Zhang, “Research on Intrusion Detection Based on Campus Network”, Intelligent Information Technology Application, 2009 Vol 1, Pages 468-471.

[7]W. Arbaugh. Guest Editor’s Introduction: Wired on Wireless. IEEE Security & Privacy, 2(3):26–27, May-Jun. 2004.

[8]Santhosh Kumar, S.; Vignesh, J.; Rangarajan, L.R.; Narayanan, V.S.; Rangarajan, K.M.; Venkatkrishna, A.L., “A Fast Time Scale Genetic Algorithm based Image Segmentation using Cellular Neural Networks”, Signal Processing and Communications, 2007. ICSPC 2007, IEEE, Pages 908-911.

[9] X. Chen, K. Makki, K. Yen, and N. Pissinou. Network Security: A Survey. IEEE Communications Surveys & Tutorials, 11(2):52–73, Second Quarter 2009. [6] N. Ye, X. Li, Q. Chen S. M. Emran, and M. Xu,
“Probabilistic techniques for intrusion detection based on computer audit data,” IEEE Trans. SMC-A, vol. 31, pp. 266–274, Jul. 2001.

[10] Baoyi Wang; Feng Li; Shaomin Zhang, “Research on Intrusion Detection Based on Campus Network”,
Intelligent Information Technology Application, 2009 Vol1, Pages 468-471.

[11] Akyazi, U; Uyar and A.S.E., “Distributed Intrusion Detection Using Mobile agenets against DDoS attacks”,Computer and Information Sciences 2008, Pages 1-6.

[12] Fiskiran, A.M; Lee, R.B., “Runtime Execution Monitoring to detect and prevent malicious code execution”, Computer
Design: VLSI in computers and Processors 2004, IEEE International Conference Pages 452-457.

[13] Jiang, M.; Munawar, M.; Reidemeister, T.; Ward, P,“Efficient Fault Detection and Diagnosis in complex Software Systems with Information- Theoretic Monitoring”, IEEE transactions on Dependable and Secure
Computing 2011, Issue 99.

[14] Nicoletta Dessì and Barbara Pes, “An Evolutionary Method for Combining Different Feature Selection Criteria in Microarray Data Classification”, Journal of Artificial
Evolution and applications, Vol 2009.

[15] Burney S. M. Aqil, Sadiq Ali Khan and Jawed Naseem, 2010, “Efficient Probabilistic Classification Methods for NIDS”, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 8, No. pp168-172