IJCSE International Journal of Computer Sciences and Engineering Open Access

Review Paper Volume-4, Issue-2 E-ISSN: 2347-2693

An Implementation of Intrusion Detection System Based on Genetic Algorithm

Kamlesh Patel[1*] and Prabhakar Sharma[2]

[1*,2] Department of Computer Science and Engineering,
Raipur Institute of Technology, Raipur (C.G.)

www.ijcseonline.org

Received: Jan /21/2016 Revised: Feb/04/2016 Accepted: Feb/19/2016 Accepted: Feb/29/2016

ABSTRACT-The intrusion detection downside is turning into a difficult task attributable to the proliferation of heterogeneous networks since the raised property of systems provides larger access to outsiders and makes it easier for intruders to avoid identification. Intrusion observation systems are accustomed detect unauthorized access to a system. By This paper I am going to present a survey on intrusion detection techniques that use genetic rule approach. Currently Intrusion Detection System (IDS) that is outlined as an answer of system security is used to spot the abnormal activities during a system or network. To this point completely different approaches are utilized in intrusion detections, however regrettably any of the systems isn't entirely ideal. Hence, the hunt of improved technique goes on. During this progression, here I even have designed AN Intrusion Detection System (IDS), by applying genetic rule (GA) to expeditiously observe numerous styles of the intrusive activities among a network. The experiments and evaluations of the planned intrusion detection system are performed with the NSL KDD intrusion detection benchmark dataset. The experimental results clearly show that the planned system achieved higher accuracy rate in distinctive whether or not the records are traditional or abnormal ones and obtained cheap detection rate.

KEYWORDS: *Intrusion Detection, Genetic Algorithm, NSL KDD, MATLAB*

I. Introduction

The Internet and native space networks area unit expanding at a tremendous rate in recent years. Whereas we tend to are getting benefits by the convenience that the latest technology has brought U.S., pc systems area unit exposed to increasing security related threats that generated externally or internally. Completely different however complementary technologies are developed and deployed to protect organizations' pc systems against network attacks, as an example, anti-virus software package, firewall, message coding, secured network protocols, countersign protection, and so on. Despite completely different protection mechanisms, it's nearly not possible to own a totally secured system [1]. Therefore, intrusion detection is changing into a progressively necessary technology that monitors network traffic and identifies network intrusions like abnormal network behaviours, unauthorized network access, and malicious attacks to pc systems [2] .There are a unit 2 general classes of intrusion detection systems (IDSs): misuse detection and anomaly detection Misuse sight ion systems detect intruders with legendary patterns, and anomaly detection systems establish deviations from traditional behaviours of networks and alert for dangerous unknown attacks. Some IDSs integrate each misuse and anomaly detection and kind hybrid detection systems. The IDSs also can be classified into 2 classes betting on wherever they appear for intrusions. A host-based IDS monitors activities related to a specific host, and a network-based IDS listens to network traffic. Variety of sentimental computing primarily based approaches are planned for detective work network intrusions . Soft computing refers to a gaggle of techniques that exploit the tolerance for inexactness, uncertainty, partial truth, and approximation to realize lustiness and low answer value. The principle constituents of sentimental computing area unit mathematical logic (FL), Artificial Neural Networks (ANNs), Probabilistic Reasoning (PR), and Genetic Algorithms .When we are going to use for intrusion detection, the soft computing techniques area unit typically utilized in conjunction with rule-based professional systems effort professional information , wherever the information is diagrammatical as a group of if-then rules. Despite completely different soft computing primarily based approaches having been planned, the probabilities of mistreatment the techniques for intrusion detection area unit still under-utilized. In this paper, we tend to give a Genetic algorithm based approach to network misuse detection. GA is chosen as a result of a number of its nice properties, e.g., sturdy to noise, no gradient info is needed to search out a worldwide optimum or sub-optimal answer, self-learning capabilities, etc. mistreatment GAs for network intrusion detection has tried to be a cheap approach. During this work, we tend to implement a software package supported the given approach. Now a day, security drawback becomes a serious issue thanks to great deal of use of web and ADP system. Any network attacks on a system violets integrity, confidentiality, and convenience. To decrease such AN influence on a network we want intrusion detection system. There are a unit varied varieties of intrusion detection system like host primarily based IDS, Network primarily based IDS. The Host primarily based IDS run on singly on system. The Network primarily based IDS monitors' traffic on a network for any suspicious activity.

Attacks types-Intrusion connected information is loosely classified in four different types of attacks [4] [6] as explained below:

1) Dos: (Denial of service): may be a category of attack wherever associate offender makes a computing code section or memory resources very busy or too full to handle legitimate request, so denying legitimate users access to a machine.

2) R2L (Unauthorized access from a distant machine): A remote to user attack may be a category wherever associate offender sends packet to a machine over a network, then exploits the machine's vulnerability to lawlessly gain native access as a user.

3) U2R (Unauthorized access to native super user (Root):U2R exploits area unit a category of attack wherever offender begin out with access to a traditional user account on the system and is in a position to use vulnerability to achieve root access to the system.

4) inquisitor (Surveillance associated different probing): Is a category of attack wherever an offender scans a network to assemble info or realize better-known vulnerability .An offender with a map of machines and services that area unit out there on a network will use the knowledge to appear for exploits.

(ii) IDS types-There are a unit varied styles of IDS supported its utilization, these are

(i) Host based mostly Intrusion Detection System (HIDS): Get audit information from host audit trails. Detect attacks against one host.

(ii) Network based mostly Intrusion Detection System (NIDS): Use network traffic because the audit information supply, relieving the burden on the hosts that sometimes offer traditional computing services, discover attacks from network.

(iii) Distributed Intrusion Detection System (DIDS): Gather audit information from multiple host and probably the network that connects the hosts, discover attacks involving multiple hosts.

II. Previous Genetic analysis for Network Intrusion Detection

Wei Li [8] wrote a proposal for mistreatment GA in an exceedingly NIDS and Ren Hui Gong [Gong] followed along with his implementation. Li set the muse for making a system mistreatment Genetic Algorithms that analyses office information sets, and Gong created a projected implementation mistreatment ECJ [Ecolab] (A Java-based organic process Computation analysis System). Gong provided pseudo code and sophistication diagrams (one aware of the ECJ library might most likely implement the algorithm). Li projected mistreatment office information sets [DARPA] from university Lincoln Laboratory for coaching and testing.

In each Li's proposal and Gong's approach they produce a fitness operate and a body sort for the Genetic Algorithms.

III. Software Tools

To implement entire experimental work for planned analysis set up , several data processing tools is utilized ,these tools provides associate degree integrated graphical programmer (GUI) facility to construct models on the premise of the many data processing based mostly algorithms out there ,one will integrate quite one techniques along to develop ensemble model .[5] [9] Tools conjointly provides some in-built feature choice techniques .The detail of varied tolls square measure explores in additional detail as below:

(a) WEKA (Waikato Environment for Knowledge Analysis): Is a well-liked suite of machine learning package written in Java, benefits of wood hen include:

- Free availableness below the antelope General Public License
- portability, since it's totally enforced within the Java artificial language and so runs on virtually any trendy computing platform
- A comprehensive assortment of information preprocessing and modeling techniques
- ease of use as a result of its graphical user interfaces

(b) CLEMENTINE: Clementine is wide considered the leading data processing bench as a result of it delivers the most come on investment within the minimum quantity of your time. Not like different data processing workbenches, that fail to actually support the whole business method of information mining and focus just on models for enhancing performance. Clementine supports the whole {data mining data methoding} process to shorten time-to-solution.

(c) MATLAB: MATLB may be a problem-oriented language and interactive atmosphere for numerical computation, visual image, and programming. Victimization MATLAB, you'll be able to analyze information, develop algorithms, and build models and applications. The language, tools, and constitutional scientific discipline functions change you to explore multiple approaches and reach an answer quicker than with spreadsheets or ancient programming languages, like C/C++ or Java.

IV. Machine Learning Techniques-

There are numerous machine learning techniques associated with data processing and applied mathematics technique as explained below:

(a) Data processing techniques:

Decision tree primarily based techniques are wide used for classification of intrusion connected knowledge as a result of its capability of exploring data in terms of straightforward if-then rules ,a decision tree is initial inducted supported coaching knowledge so tested exploitation testing samples[13] [10]. Numerous call tree primarily based techniques to be accustomed style framework for IDS are explained in additional detail as explained below:

(i) Classification and Regression Technique (CART)
(ii) Chi-Squared Automation Interaction Detection (CHAID)
(iii) Iterative Dichotomize three (ID 3)

(b) Soft computing techniques:

Soft computing differs from conventional computing, the model for soft computing is the human mind. During this the principal subsequent are tools, techniques of soppy computing are fuzzy logic (FL), neuralnetwork (NN), support vector machine (SVM), evolutionarycomputation (EC).

(i) Artificial Neural Network
(ii) Hybrid soft computing

(c) Applied mathematics Techniques:

Statistical models (David L. Olson et al., 2008) involving a latent structure usually support cluster, classification, and different data processing tasks. due to their ability to cope with lowest data and droning labels during a systematic fashion, applied mathematics models of this kind have recently gained quality, and success stories are often found during a form of applications; for instance, population genetic science, scientific publications, words and pictures, incapacity analysis, fraud detection, biological sequences &amp; networks. There are some applied mathematics model describe below:

(I)Bayesian web
(II)Support Vector Machine (SVM)

(d) Ensemble Model:An ensemble model may be a combination of 2 or a lot of models to avoid the drawbacks of individual models and to attain high accuracy. Sacking and boosting (Han, J., &amp; Michelin, K., 2006) are 2 techniques that use a mixture of models. every combines a series of k learned models (classifiers), M1, M2...Mk, with the aim of making Associate in Nursing improved composite model, M. each sacking and boosting are often used for classification.

V. Genetic Algorithm

The Genetic algorithms employ metaphor which is fromgenetics and biology to iteratively evolve a population of initial Individuals to the population of very high quality individuals, where each individual represents a solution of the Problem to be solved and is made of a fixed number of genes [7]. The number of possible values of each gene is called the cardinality of the gene.

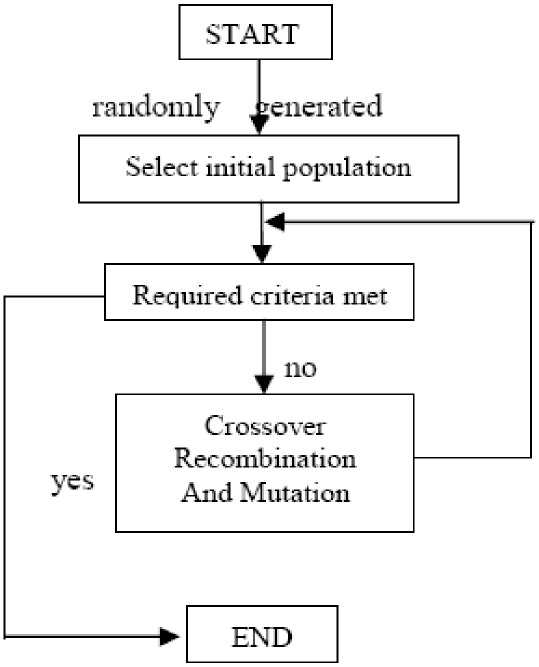

Fig. Structure and Function of Genetic Algorithm

Functioning of genetic algorithmic rule starts with every which way generated population of people. Through if generations varied, these population evolved and individuals' quality gets improved. In each generation, 3 basic operators of genetic algorithmic rule i.e. selection, crossover and mutation area unit applied to every individual. Crossover suggests that exchanging the factors between 2 bodies' whereas mutation suggests that random dynamical of a worth of a every which way chosen gene of a chromosome. These people area unit illustration of the matter needed to be resolved. Completely different positions of every individual are often encoded as bits, characters and numbers. Here, the numbers of best-fit people area unit designated. For this user outlined fitness operate is employed. Fitness operate is employed to live quality of everybody. Remaining people area unit paired and through method of crossover, new offspring is made by part exchanging their genes.

When genetic algorithmic rule is employed for drawback resolution, 3 factors can have impact on the effectiveness of the algorithmic rule, they Are:-

a. the choice of fitness operate

b. The illustration of people and

c. The values of the genetic parameters.

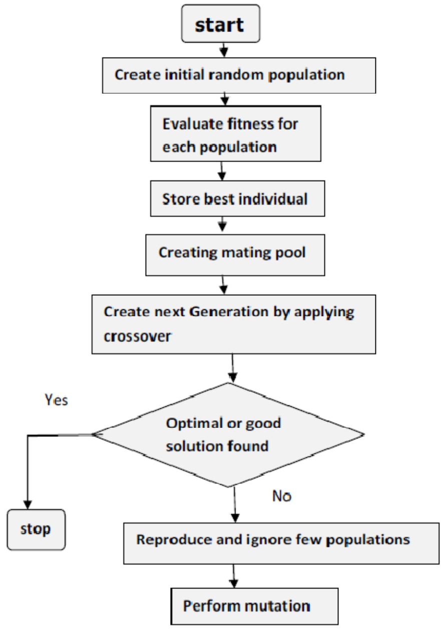
start
Create initial random population
Evaluate fitness for each population
Store best individual
Creating mating pool
Create next Generation by applying crossover
Yes
Optimal or good solution found
No
stop
Reproduce and ignore few populations
Perform mutation

Figure. Flowchart (Processing steps of GA.)

Genetic Algorithms square measure a looking out algorithm designed to mimic the manner nature reproduces itself and betters itself in doing therefore. [3] Genetic algorithms have many people in its population, every one of that might be a possible answer to a retardant. Each one of these people would be following path to a doable answer, which implies that it is even doable for the search to seek out quite one solution. Completely different researchers have created many Genetic Algorithms, and every one of them square measure very different from one another. They all, however, show the characteristics of the genetic algorithmic program that follow these basic steps:

1. Randomly produce a population of individuals.
2. Is to judge the population to envision once the new population has been created, it is time to change the people so they are different from the opposite generation. There square measure two possible operations, which may be performed here, crossover and mutation.

1. Crossover

This occurs once 2 people that have been paired off, exchange chromosomes. A random number is generated. Range theamount the quantity should be between 1 and also the 1(is the most number of bits within the bit string). The bit string is taken into account to be the individual and also the bits square measure the chromosomes. Once the random range has been generated, then all digits at that time position within the bit string are exchanged. That people can contribute to the nextGeneration? Step 3 is to change the new generation of individuals once they need been paired off. The final step is to discard the previous population and perform step 2 on the new population. Once step 3, above, has been completed, the algorithm jumps back to step 2. The loop will only stop once one among the people has been evaluated and is claimed to be either terribly near the solution, or it's found the answer.

2) Mutation

Another amendment a personal might undergo is known as mutation [11] [12]. Not like crossover, it only involves one individual. Counting on what algorithm it's (there square measure several genetic algorithms), mutation isn't terribly probably to occur. Usually, the possibility of a mutation occurring one} of the chromosomes is ready to 1 in thousands. There square measure many alternative genetic algorithms, some of that don't incorporate mutation, and when they do, they handle mutation in a very completely different manner to therest. The fundamental example may have the computer system generate a random range that decides whether or not slightly is to endure mutation. Once slightly is chosen for mutation, some algorithms merely that don't incorporate mutation, and after they do, they handle mutation in a special thanks to the rest! The fundamental example may have the pc generate a random range that decides whether or not or not slightly is to undergomutation. Once slightly is chosen for mutation, some algorithms merely

VI. BENEFITS OF USING GENETICALGORITHM FOR INTRUSION DETECTIONARE

Genetic algorithms area unit per se parallel. Because of multiple offspring, they'll explore the solution house in multiple directions directly. Correspondence permits genetic rule to implicitly assess several schemas directly. This make them well matched to resolution issues where space of potential resolution is actually large. Genetic rule primarily based systems are often retrained easily. This improves its risk to feature new rules and evolve intrusion detection system.

VII. CONCLUSION

In this paper, we present and implemented an Intrusion Detection System by applying genetic algorithm that will efficiently detect various types of network intrusions and malicious activities. To implement and measure the performance of our system we used the standard NSL-KDD benchmark dataset and obtained reasonable detection rate. In near future we will try to enhance our intrusion detection system by using more statistical analysis and with better and may be more complex equations.

VIII. ACKNOWLEDGEMENT

I would like to extend my sincere thanks and gratefulness to my Guide Mr. Prabhakar Sharma sir , Assistant Professor ,Department of Computer Science and Engineering, Raipur Institute of Technology Raipur, (C.S.V.T.U. University, Bhilai), India for his such kind help and guiding me for preparing this article.

IX. REFRENCES

[1] R.Elamaran and R.Mala, "A Study on Network Intrusion Detection System (NIDSs) In Virtual Network Structure", International Journal of Computer Sciences and Engineering (IJCSE), Vol. 03, Issue - 11,**November-2015**, pp.**59 – 164**.

[2] Mostaque Md. Morshedur Hassan, LCB College, Maligaon, Guwahati, Assam, India,"Current Studies on Intrusion Detection System, Genetic Algorithm and Fuzzy Logic", International Journal of Distributed and Parallel Systems, Vol. 4, Issue-2, **March-2013**,pp.**35-47**.

[3] Y.Dhanalakshmi and Dr. I. Ramesh Babu, "Intrusion Detection Using Data Mining Along Fuzzy Logic and Genetic Algorithms", Dept. of Computer Science & Engineering , Acharya nagarjuna University, Guntur, A.P. India International Journal of Computer Science & Network Security (IJCSNS), Vol.8,Issue-2,**Februry-2008**,pp.**27-32**.

[4] Zorana Banković, José M. Moya, Álvaro Araujo, Slobodan Bojanić and Octavio Nieto-Taladriz, "A Genetic Algorithm based Solution for Intrusion Detection", Journal of Information Assurance and Security(JIAS), Vol.4 , Issue-3 ,**June-2009**, pp.**192-199**.

[5] Shelly Xiaonan Wu, Wolfgang Banzhaf, "The use of computational intelligence in intrusion detection systems: a review", Applied Soft Computing, Vol.10, Issue-01, **January-2010**, pp.**1–35**.

[6] Mohammad Sazzadul Hoque, Md. Abdul Mukit & Md. Abu Naser Bikas, "An Implementation of Intrusion Detection System using Genetic Algorithm" , Department of Computer Science and Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh, International Journal of Network Security and Its Applications (IJNSA),Vol.4,Issue-2,**March-2012**, pp.**109-120**.

[7] S Selvakani Kandeeban, and Rengan S Rajesh , Department of Computer Applications, Jaya Engineering College1 Chennai, Tamilnadu, 602 024, India. Dept. of CSE, MS University, Tirunelveli, Tamilnadu, 627 009, India,"Integrated Intrusion Detection System using Soft computing", International Journal of Network Security, Vol.10, Issue-2, **March -2010**,pp.**87-92.**

[8] W. Lu and I. Traore, Department of Electrical and Computer Engineering, University of Victoria, Victoria B.C., Canada "Detecting New Forms of Network Intrusion Using Genetic Programming", *Computational Intelligence,* vol. 20, Issue-03, **August -2004**, pp.**475-494**.

[9] K. Burbeck & N.Y. Simmin (2007), Department of Computer and Information Science, Linkoping University ,Sweden,"Adaptive Real-Time Anomaly Detection with Incremental Clustering", Information Security Technical Report, Vol. 12, Issue- 1 ,**07-March- 2007**,pp.**56–67**.

[10] T.S. Chou, K.K. Yen & J. Luo , "Network Intrusion Detection Design using Feature Selection of Soft Computing Paradigms", International Journal of Computational Intelligence, Vol. 4, Issue-3, **2008,** pp.**196–208**.

[11] Bhavani M. Thuraisingham,Latifur Khan, Mamoun Awad , University of Texas at Dallas, Dallas , USA,"A New Intrusion Detection System using Support Vector Machines and Hierarchical Clustering", The International Journal on Very Large Data Bases, Vol.16, Issue-04,**October-2007**,pp.**507–521**.

[12] S. M. Aqil Burney,M. Sadiq Ali Khan and Jawed Naseem,Department of Computer Science ,University of Karachi , Pakistan, "Efficient Probabilistic Classification Methods for NIDS", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 8 ,Issue-08 , **November-2010**, pp.**168-172**.

[13] Baoyi Wang; Feng Li; Shaomin Zhang, "Research onIntrusion Detection Based on Campus Network", Intelligent Information Technology Application, Vol.01, **2009**, pp.**468-471**.