INTRUSION DETECTION SYSTEM USING ARTIFICIAL INTELLIGENCE TECHNIQUE-GENETIC ALGORITHM

[1]**Priyanka A. Motekar,**
M.E (C.S.E.),Scholar, Pankaj Laddhad Institute of Technology and Management Studies, Buldhana-443001, Maharashtra

[2]**Dr. Pradip M. Jawandhiya,**
Principal Pankaj Laddhad Institute of Technology and Management Studies, Buldhana-443001, Maharashtra.

Abstract—- With the fast growth of net in recent years, laptop systems face enlarged variety of security threats. Despite various technological innovations for info assurance, it's still terribly troublesome to shield laptop systems. Therefore, unwanted intrusions happen once the particular software package systems square measure running. Completely different soft computing primarily based approaches are planned to notice network attacks. This paper presents varied approaches to network intrusion detection like genetic algorithm (GA), associate increased deciding by rule-list i.e. fuzzy classifier and artificial neural network classifier primarily based approach to network intrusion detection. The project additionally shows the potency of algorithms in terms of your time for classification. These classification rules square measure accustomed notice networking attacks or intrusions. The planned system is applied on KDDCup99 Dataset to yield a lot of economical and effective classification rules.

Index Terms—: KDDCup99 Dataset, Genetic Algorithm (GA), Fuzzy Classifier, IDS, DOS Attack

I. INTRODUCTION

The number of intrusions into laptop systems is growing as a result of new automatic hacking tools square measure showing each day, and these tools together with varied system vulnerability info square measure simply obtainable on the net. The matter of intrusion detection has been studied extensively in laptop security and has received plenty of attention in machine learning and data processing. Despite increasing awareness of network security, the prevailing solutions stay incapable of totally protective web applications and laptop networks against the threats from ever-advancing cyber-attack techniques like DoS attack and laptop malware. Developing effective and adaptive security approaches, therefore, has become a lot of vital than ever before. the normal security techniques, because the initial line of security defense, like user authentication, firewall and encryption, square measure poor to completely cowl the complete landscape of network security whereas facing challenges from ever-evolving intrusion skills and techniques thence, another line of security defense is extremely counseled, like Intrusion Detection System (IDS).

Redundant and impertinent options in information have caused a long-run downside in network traffic classification. These options not solely curtail the method of classification however conjointly forestall a classifier from creating correct selections, particularly once dealing with massive information. During this paper, we have a tendency to propose a mutual info primarily based formula that analytically selects the optimum feature for classification i.e. information reduction technique named. This mutual info primarily based feature choice formula will handle linearly and nonlinearly dependent information options. Its effectiveness is evaluated within the cases of network intrusion detection. Associate degree Intrusion Detection System (IDS), with 3 completely different classification techniques has been projected named genetic formula (GA), associate degree increased higher cognitive process by rule-list i.e. fuzzy classifier and artificial neural network classifier primarily based approach to network intrusion detection. The performance of evaluated exploitation intrusion detection analysis dataset, specifically KDD Cup ninety nine dataset. The analysis results show that our projected approaches for IDS to realize higher accuracy and lower machine value compared with the progressive ways with relevancy time for classification to observe the malicious information within the KDD Cup ninety nine datasets.

Figure 1

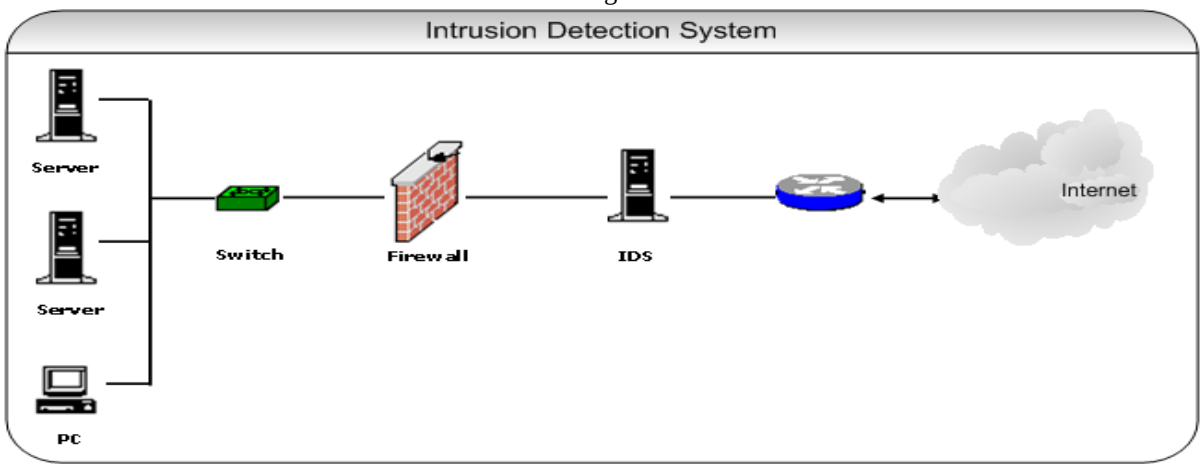


Fig 1. Typical Intrusion Detection System

II. LITERATURE SURVEY

1. Paper Name: Using Genetic Algorithm for Network Intrusion Detection
   Authors: Wei Li
   Description: This paper describes a method of applying Genetic algorithmic program (GA) to network Intrusion Detection Systems (IDSs). A quick summary of the Intrusion Detection System, genetic algorithmic program, and connected detection techniques is conferred. Parameters and evolution method for GA square measure mentioned well. Not like different implementations of a similar downside, this implementation considers each temporal and abstraction data of network associations in cryptography the network connection data into rules in IDS.
2. Paper Name: Evolving Fuzzy Classifiers for Intrusion Detection
   Authors: Jonatan Gomez and Dipankar Dasgupta.
   Description: The normal and therefore the abnormal behaviors in networked computers area unit exhausting to predict, because the boundaries cannot be well outlined. This prediction method sometimes generates false alarms in several anomaly based mostly intrusion detection systems. However, with mathematical logic, the warning rate in determinant intrusive activities is reduced, wherever a collection of fuzzy rules is employed to outline the conventional and abnormal behavior during an electronic network, and a fuzzy illation engine is applied over such rules to work out intrusions. This paper presents some results and reports the performance of generated fuzzy rules in classifying differing kinds of intrusions.
3. Paper Name: Fitness Function for Genetic Algorithm used in Intrusion Detection System
   Authors: Firas Alabsi and Reyadh Naoum
   Description:Computer network usage inflated apace at the last decades, the trespassers tried to satisfy their desires by many varieties of attack counting on the intruder objectives, this encourage the researchers to search out additional and additional solutions to sight those attacks. Intrusion Detection System won't to sight the attack. Genetic rule won't to support IDS. This paper can show the Fitness operate, discuss it and compare it with another Fitness operate to envision its validity.
4. Paper Name: Building an intrusion detection system using a filter-based feature selection algorithm
   Authors: Mohammed A. Ambusaidi, Xiangjian He, Priyadarsi Nanda and Zhiyuan Tan.
   Description:Redundant and impertinent options in information have caused a long term drawback in network traffic classification. These options not solely hamper the method of classification however conjointly forestall a classifier from creating correct selections, particularly once managing massive information. Associate Intrusion Detection System (IDS), named Least sq. Support Vector Machine primarily based IDS (LSSVM-IDS), is constructed exploitation the options elite by our projected feature choice algorithmic rule.
   Current network traffic data, which are often huge in size, present a major challenge to IDSs. These "big data" slow down the entire detection process and may lead to unsatisfactory classification accuracy due to the computational difficulties in handling such data. Classifying a huge amount of data usually causes many mathematical difficulties which then lead to higher computational complexity. There are approaches such as artificial neural network classifier and fuzzy classifier techniques in the literature.

- **Disadvantages of Existing System:**
  1. Redundant and impertinent options in information have caused a long-term drawback in network traffic classification.
  2. These options not slow down the method of classification however conjointly forestall a classifier from creating correct selections, particularly once managing huge information.
  3. Low performance.

III. PROPOSED SYSTEM

We propose a virus detection system placed at the network egress points to detect malware infection which relies on DNS to locate malicious threats in the network. Implantation of IDS system for network dataset named KDD Cup 99 with genetic algorithm for classification.

We conduct complete experiments on two well-known IDS datasets additionally to the dataset used. This can be important in evaluating the performance of IDS since KDD dataset is noncurrent and doesn't contain most novel attack patterns in it. Additionally, these datasets are often utilized in the literature to evaluate the performance of IDS.

Different from the detection framework planned that styles just for binary classification; we have a tendency to style our planned framework to think about multiclass classification issues. This can be to indicate the effectiveness and therefore the feasibleness of the planned technique.

- **Characteristics:**

A. Instance selection

Instance selection is a technique to reduce the number of instances by removing noisy and redundant instances. By using this technique original data sets are reduced by removing non-representative instances. For a given data set in a certain application, instance selection is to obtain bug reports in bug data

B. Feature selection

Feature selection aims to obtain a subset of relevant features. It is a preprocessing technique used for selecting a reduced set of features for large Scale data sets. In our work we leverage the combination of Instance selection and Feature selection to generate a data set.

- **Advantages of Proposed System:**
  1. Recently, IDS alongside with anti-virus software has become an important complement to the security infrastructure of most organizations.
  2. IDS is designed to detect malicious domains used for crafted malware in APT attacks and to detect infected machines.
  3. High performance and less complexity.

IV. GENETIC ALGORITHM

1. **[Start]** Generate random training dataset of *n* attributes (suitable solutions for the problem)
2. **[Fitness value]** Evaluate the fitness *f(x)* of each attribute *x* in the population
   Here fitness value = identified malicious attributes in the dataset.
3. **[New population]** Create a new population by repeating following steps until the new population is complete.
   1. **[Selection]** Select parent attribute i.e. dataset from a training set according to their fitness (the better fitness, the bigger chance to be selected)
   2. **[Crossover]** With a crossover probability cross over the parents to form new offspring (children) i.e. new attribute which is has same fitness value to the existing fitness values in the dataset. If no crossover was performed, then attribute is the exact copy of parents.
   3. **[Accepting]** Place new offspring (attribute) in the new population (dataset).
4. **[Replace]** Use new generated population for a further run of the algorithm.
5. **[Test]** If the end condition is satisfied, **stop**, and return the best solution in current population
6. **[Loop]** Go to step **2**

V. MATHEMATICAL MODEL

Let w is the set of whole system which consist
W= {input, process, output}

Input:-
File= {f1, f2,.....,fn}

Process:-
In this paper we have used Genetic algorithm.
This algorithm filters the features or attributes & provides the exact results to query.

Output:-Accurate Results.

VI. SYSTEM ARCHITECTURE

Figure 2

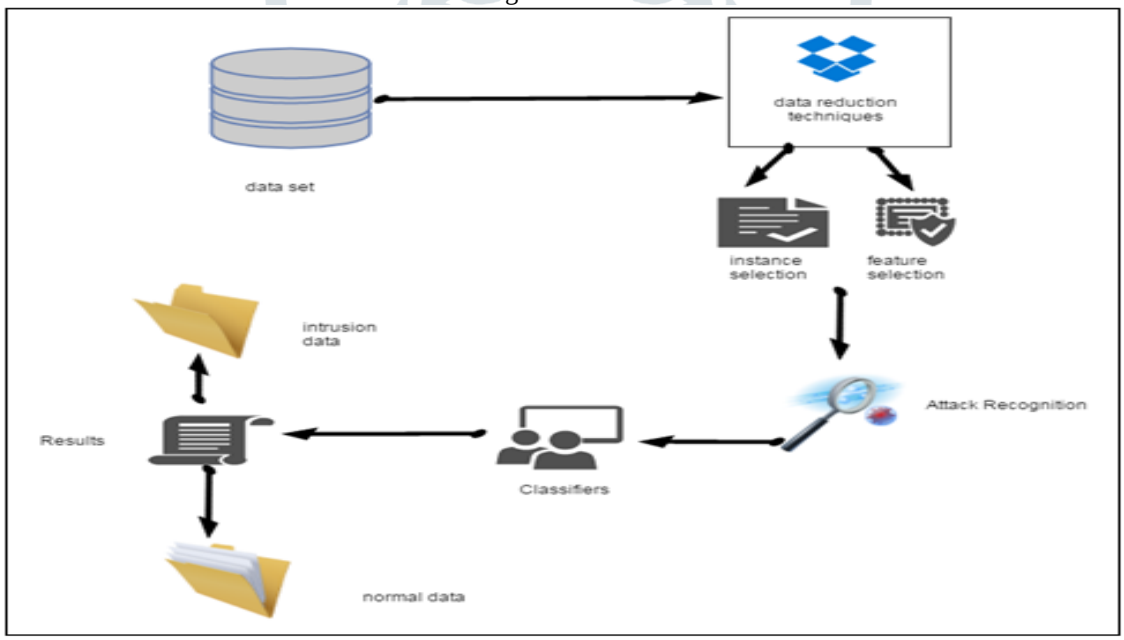
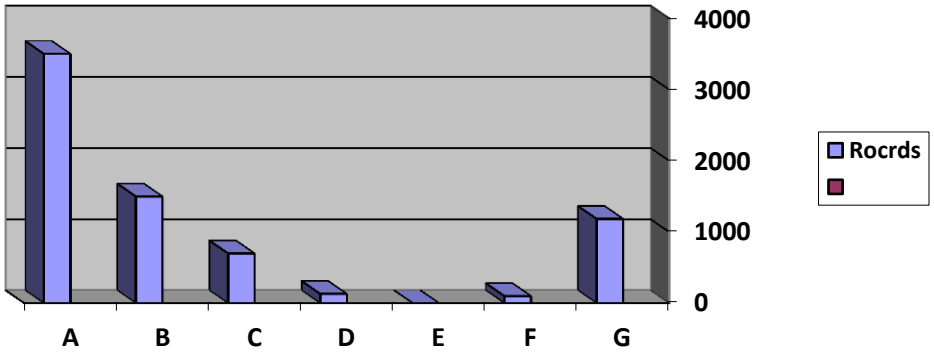


Fig 2. Proposed System Architecture

VII. RESULT ANALYSIS

	A	B	C	D	E	F	G
Records	3500	1497	698	131	1	99	1183

Where,
A=Total Records
B=Malicious
C=DOS
E=Probe
F=UR2
G=R2L
H=Normal



VIII. CONCLUSION AND FUTURE SCOPE

In this, a proposed a system IDS placed at the network egress points to detect malware infections inside the network combined network traffic analysis. Proposed an approach for network intrusion detection for KDD Cup 99 dataset with genetic algorithm (GA). The experimental results show that this security approach is feasible for improving the sustainability of the system and is good at detecting malware infections attacks. We also evaluated the efficiency of proposed Genetic Algorithm (GA) with respect to time required for different classification techniques in the literature. In near future we will try to improve our intrusion detection system with the help of more statistical analysis and with better and may be more complex equations with the help new classification techniques and also Combining knowledge from different security sensors into a standard rule base is one field of research.

IX. ACKNOWLEDGMENT

I would like to thank to my guide **Dr. P. M. Jawandhiya**, for his help and support during my project work. He gave me the knowledge from his own experience. His guidance helped me in growing my confidence which would help me in my future. So I am sincerely thankful to my guide.

I express thanks and gratitude to Head of Department and Principal **Dr. P. M. Jawandhiya** for his encouraging support and co-operation in this paper and sincere thanks and respect to him for giving his valuable guidance inspiration and affectionate encouragement to embark this paper.

My thanks also goes to other staff members of Computer Science & Engineering Department, my parents and my friends who are directly or indirectly involved in my work. I am extremely grateful to all the helping contribution for the successful completion of this project.

REFERENCES

[1] ChSatyaKeerthi N.V.L., Prasanna P.I., Priscilla B.M., "Instuction Detection system Using Genetic Algorithm",Int. Journal of P2P Network rends and Technology, vol.1.no. 2.pp 1-7, 2011.

[2] Janusz Starzyk ,Jing Pang," Evolvable Binary Artificial Neural Network for Data Classification", Evolutionary Computation pp 5576-5783(2000)

[3] Guoqiang Peter Zhang, "Neural Networks for Classification: A Survey", IEEE Transaction on systems management and Cybernatics-Applications and Reviews Vol-30, No-4,pp 451-462(2000).

[4] Jiang M., Munavar M., Reidemeister T., Ward P.,"Efficient Fault Detection and Diagnosis in Complex Software Systems with Information- Theoretic Monitoring"IEEE Trans.On Dependable and Secure Computing.Issue 99, 2011.

[5] Chittur A.,""Model Generation for an Intrusion Detection System Using Genetic Algorithms", 2011.

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

[7] Pedro A. Diaz-Gomez and Dean F. Hougen "Three Approaches to Intrusion Detection Analysis And Enhancements", National Computer And Information Security Conference Acis2006 .

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

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

[10] Alabsi,F., Naoum,R., "Fitness Function for Genetic Algorithm used in Intrusion Detection System",International Journal of Applied Science And Technology.Vol. 2.no 4, 2012.

[11] Kandeeban,S.S., Rajesh R.S.,"A Mutual Construction For IDS Using GA", Int. Journal of Advance Science And Technology,vol.29, 2011.

[12] Uppalaiah B., Anand K., Narsimha B., warajS., BharatT., "Genetic Algorithm Approach to Intrusion Detection System", IJCST vol.3.1,2012.

[13] Owais S.S.J., Kromer P., Snasel V., "Implementing GP onoptimizing Boolean and Extended Boolean Queries in IRs withRespectto Users Profiles", Proc. IEEE lCCES'06 Egypt. pp412-417.2006

[14] Goyal A., Kumar C., "GA-NIDS: A Genetic Algorithm based Network Instruction Detection System", 2008

[15] Mohammad S. H., MukitMd.A.,BikasMd.A. N.," An Implementation of Intrusion Detection System Using Genetic Algorithm", Int.Journal of Network Security and Its Applications, vol.4 no.2. pp 109-119, 2

[16] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, W.-Y. Lin, Intrusion detection by machine learning: A review, Expert Systems with Applications 36 (10) (2009) 11994–12000.