IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 4, Ver. IV (July – Aug. 2015), PP 65-68
www.iosrjournals.org

A Study on A Hybrid Approach of Genetic Algorithm & Fuzzy To Improve Anomaly or Intrusion

Er.Kamaldeep Kaur[1], Er. Simranjit Kaur Dhindsa[2]
[1](CSE, ACET/PTU ,India)
[2](CSE, ACET/PTU ,India)

Abstract: *This paper describes a technique of applying Genetic Algorithm (GA) and fuzzy to network Intrusion Detection Systems (IDSs). A brief overview of a hybrid approach of genetic algorithm and fuzzy to improve anomaly or intrusion is presented. . This paper proposes genetic algorithm and fuzzy to generate that are able to detect anomalies and some specific intrusions. The goal of intrusion detection is to monitor network activities automatically, detect malicious attacks and to establish a proper architecture of the computer network security. Experimental results demonstrate that we can achieve better running time and accuracy with these modifications.*

Keywords: *anomaly***,** *clustering, fuzzy algorithm, fuzzy set***.,** *genetic algorithm, intrusion detection.*

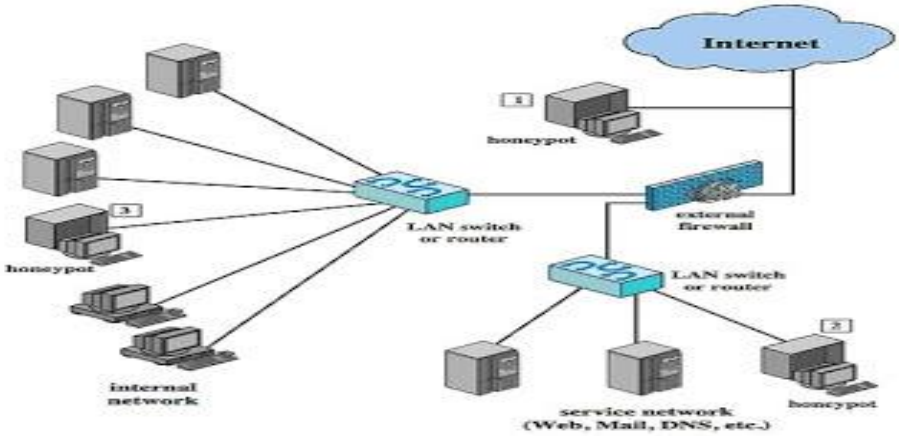
I. Introduction

An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, and reporting attempts. In addition, organizations use IDPSs for other purposes, such as identifying problems with security policies, documenting existing threats and deterring individuals from violating security policies. IDPSs have become a necessary addition to the security infrastructure of nearly every organization.

II. Anomaly

1.Anomaly- Anomalies is the data within the database; it is the copy of the original data it needs to be updated in order to avoid problems such as viewing the website. There are four types of anomalies which are Insertion anomaly, Deletion anomaly, Duplicate entry and Modify (Update anomaly).

Database anomalies, are really just unmatched or missing information caused by limitations or flaws within a given database. Databases are designed to collect data and sort or present it in specific ways to the end user. Entering or deleting information, be it an update or a new record can cause issues if the database is limited or has 'bugs'.



1.1 Insertion anomaly- An insertion anomaly means that it is difficult to insert new records into the database.
1.2 Update anomaly- An update anomaly occurs when the same data item has to be updated more than once. This can lead to errors and inconsistency of data.
1.3 Deletion anomaly A deletion anomaly occurs when data is lost because of the deletion of other data.

III. Techniques Of Intrusion Detection

1. Clustering-Clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data. A loose definition of clustering could be "the process of organizing objects into groups whose members are similar in some way". A cluster is a collection of objects which are "similar" between them and are "dissimilar" to the objects belonging to other clusters.

1.1 Goal of Clustering-The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. But how to decide what constitutes a good clustering? It can be shown that there is no absolute "best" criterion which would be independent of the final aim of the clustering. Consequently, it is the user which must supply this criterion, in such a way that the result of the clustering will suit their needs. For instance, we could be interested in finding representatives for homogeneous groups (data reduction), in finding "natural clusters" and describe their unknown properties ("natural" data types), in finding useful and suitable groupings ("useful" data classes) or in finding unusual data objects (outlier detection).

1.2 Possible Applications- Clustering algorithms can be applied in many fields, for instance:

- **Marketing**: finding groups of customers with similar behavior given a large database of customer data containing their properties and past buying records;
- **Biology**: classification of plants and animals given their features;
- **Libraries**: book ordering;
- **Insurance**: identifying groups of motor insurance policy holders with a high average claim cost; identifying frauds;
- **City-planning**: identifying groups of houses according to their house type, value and geographical location;
- **Earthquake studies**: clustering observed earthquake epicenters to identify dangerous zones;
- **WWW**: document classification; clustering weblog data to discover groups of similar access patterns.

1.3 Fuzzy C-Means Clustering-Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. It is based on minimization of the following objective function. The membership of fuzzy variable in a fuzzy set is determined by a function that produces values within the interval(0,1). Fuzzy rules can be "if the temperature is cold and the wind is strong than wear warm clothes" where temperature and wind are antecedent fuzzy variables. Wear is a consequent fuzzy variables and cold, strong and warm clothes are fuzzy sets.

2. Genetic Algorithm-Genetic algorithm use ideas based on the language of natural genetics and biological evolution. Genetic allows humans to contribute solution suggestions to the evolutionary process. Genetic algorithm finds application in computational science, engineering, economics, chemistry, manufacturing. Genetic algorithm requires two functions-

- Genetic representation-it can encode appearance, behaviour, physical qualities of individuals. Designing a good genetic representation is expensive and evolvable is a hard problem in evolutionary computation.
- Fitness function- it is a particular type of objective function that is used to summarize as a single figure of merit. It flow close a given design solution is to achieving the set aims.

2.1 Genetic Algorithm Advantages To Intrusion Detection Systems

The implementation of genetic algorithms offers many advantages to intrusion detection systems. The benefits of using genetic algorithms for intrusion detection can be summarized as:

- Genetic algorithms offer intrusion detection systems an intrinsic parallelism.
- Genetic algorithms are capable of working in multiple directions simultaneously. This makes them beneficial for analyzing the huge volumes of multi-dimensional data to be processed by an intrusion detection system.
- Genetic algorithms work with populations of solutions rather than a single solution. This makes them suitable for behaviour based intrusion detection, where the behaviour attributes may exhibit varying values.
- Genetic algorithms are highly re-trainable. Therefore, using genetic algorithms for intrusion detection will add to the adaptability of the system.
- Genetic algorithms evolve over time by using crossover and mutation. Property of evolving over time makes them a good choice for dynamic rule generation.

IV. Results

We have to calculate the value of fit value and time and then we plot the graph with the help of clustering and Genetic algorithm.

parameter	Fit value	Time
G_num	56.9	21.59
S number	70.61	2.73
np	90.84	6.92
f	90.84	5.17
strategy	91.02	5.15
cr	91.02	5.18
threshold	90.85	5.16

Table1.

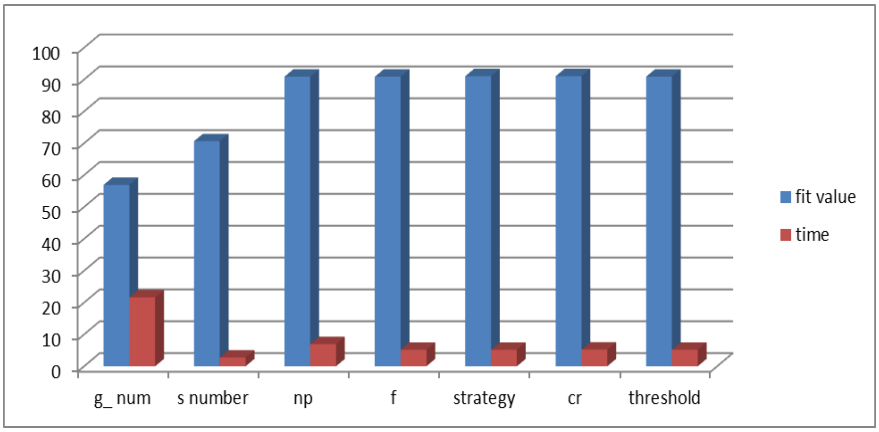


Fig.1 Anomaly Detection using Genetic Algo

parameter	Fit value	Time
G_num	35.01	39.44
S number	69.06	2.77
np	90.88	8.8
f	90.85	5.24
strategy	91.02	5.23
cr	91.02	5.26
threshold	91.02	5.37

Table 2

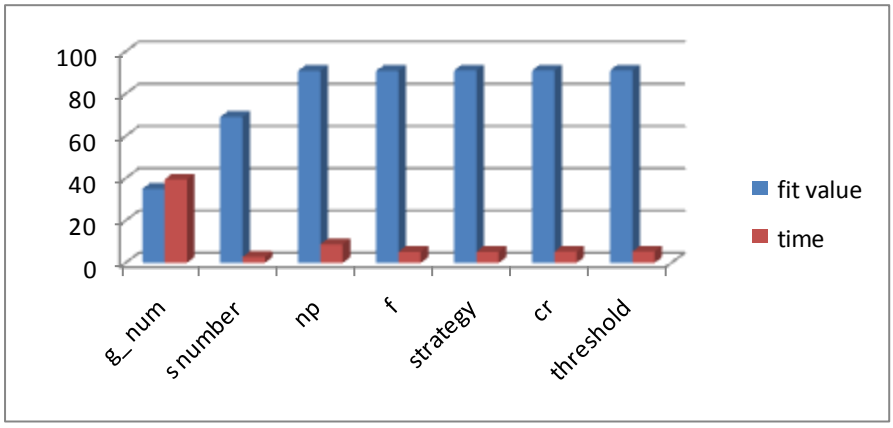


Fig.2

V. Conclusion

We have analysed an anomaly based intrusion detection system in detecting the intrusion behaviour within a network. Genetic algorithm and fuzzy decision-making module was designed to build the system more accurate for attack detection. Our experiments showed that the proposed approach works well in detecting different attacks. The accuracy of fuzzy classifiers and Genetic algorithm was good and comparable to those reported in the literature. Also, the accuracy can further be improved applying specific strategies to generate the fuzzy and genetic algorithm for each monitored parameter.

Acknowledgements

I would like to express a deep sense of gratitude to my Project Guide, Er. Simranjit Kaur Dhindsa,Assistant Professor of Department, Department of Computer Science and Engineering for her guidance and support in defining the design problem and towards the completion of my thesis work. I could learn the technique of organizing and writing quality research matter only because of her erudite teachings throughout the work. This impact has left a permanent impression on my personality and written & verbal communication. I also express my great admiration & indebtedness for the manner in which she carried out a thorough editing of our papers & the thesis, despite his overwhelming busy schedule & numerous responsibilities. Without her wise counsel and able guidance, it would have been impossible to complete the thesis in this manner.

References

[1]. Shingo Mabu, Member, IEEE, Ci Chen, Nannan Lu, Kaoru Shimada,and Kotaro Hirasawa, Member, IEEE, "An Intrusion-Detection Model Based on Fuzzy Class-Association-Rule Mining Using Genetic Network Programming" in January 2011.
[2]. Monowar H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita , "Network Anomaly Detection: Methods, Systems and Tools" in 2014.
[3]. Z. Bankovi´c, D. Stepanovi´c, S. Bojani´c, and O. Nieto-Taladriz, "Improving network security using genetic algorithm approach," Comput. Elect.Eng., vol. 33, pp. 438–451, 2007.
[4]. A. Sundaram, "An introduction to intrusion detection," Crossroads,vol. 2, no. 4, pp. 3–7, April 1996.
[5]. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection : A Survey," ACM Computing Surveys, vol. 41, no. 3, pp. 15:1–15:58,September 2009.
[6]. L. Portnoy, E. Eskin, and S. J. Stolfo, "Intrusion detection with unlabeled data using clustering," in Proc. ACM Workshop on Data Mining Applied to Security, 2001.
[7]. R. Heady, G. Luger, A. Maccabe, and M. Servilla, "The Architecture of a Network Level Intrusion Detection System," Computer Science Department, University of New Mesxico, Tech. Rep. TR-90, 1990.
[8]. A. A. Ghorbani, W. Lu, and M. Tavallaee,"Network Intrusion Detection and Prevention : Concepts and Techniques" ser. Advances in Information Security. Springer-verlag, October 28 2009.
[9]. P. SNing and S. Jajodia, "Intrusion Detection Techniques". H Bidgoli (Ed.), The Internet Encyclopedia, 2003.
[10]. F.Wikimedia,"Intrusiondetectionsystem,"http://en.wikipedia.org/wiki/Intrusion-detection system, AAFeb 2009.
[11]. H. Debar, M. Dacier, and A. Wespi, "Towards a Taxonomy of Intrusion Detection Systems," Computer Networks, vol. .31, no. 8, pp. 805–822, 1999.
[12]. D. B. P. and M. Pels, "Host-Based Intrusion Detection Systems," Faculty of Science, Informatics Institute, University of Amsterdam, Technical Report, 2005.
[13]. Aleksandar, Kumar, and Jaideep, Managing Cyber Threats: Issues, Approaches, and Challenges. Springer Science + Business Media, 2005.
[14]. D. E. Denning, "An Intrusion Detection Model," Special issue on Computer Security and Privacy, vol. 13, no. 2, pp. 222–232, 1987.