2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT)
Rajarshi Shahu College of Engineering, Pune India. Dec 2-3, 2016

Real Time Hybrid Intrusion Detection System using Signature Matching Algorithm and Fuzzy-GA

Anuja S. Desai
Dept. of Computer Engineering,
AISSMS COE Pune, Savitribai Phule
Pune University,
Pune, India
Email: anujadesai92@gmail.com

D. P. Gaikwad
Dept. of Computer Engineering,
AISSMS COE Pune, Savitribai Phule
Pune University,
Pune, India
Email: dp.g@rediffmail.com

***Abstract*- In Internet-based communication, different types of networks are used to provide services to users. Due to exploration of different types of vulnerabilities, day by day threat of attacks in network is increasing. There is possibility of getting system infected by internal or external intruders. Single computer system is shared by multiple users. Users use multiuser system by creating their own account, which is protected by unique id and password. Sometimes, authorized users attack system for some malicious purposes. Internal attacks include deleting some important files or data, altering some important data. For any intrusion detection system, it is very difficult to identify authorized internal intruders. SQL injection is one such attack, which can be launched by internal attacker. The purpose of external intruder is to perform malicious activities in remote system. To avoid these two types of attacks, the robust intrusion detection system is needed. In this paper, we have implemented hybrid intrusion detection system, which includes identification of both internal and external attacks. Signature matching algorithm is implemented to identify internal attacks. Fuzzy genetic algorithm is applied for implementation of external attacks detection. The system is hybrid and compatible in offline as well as online environment. Experimental results show that the accuracy of system is better than some of existing systems.**

Keywords— Data Mining; Insider attack; Intrusion Detection System; SQL injection; Fuzzy Genetic Algorithm;

I. INTRODUCTION

Web is generally spread in every edge of the world; personal computers all over are presented to assorted interruptions from the World Wide Web. To shield the personal computers from these unapproved assaults, viable interruption location frameworks (IDS)[1] should be utilized. Conventional occasion based learning routines for intrusion detection can just recognize known interruptions, since these strategies group examples in view of what they have realized. They scarcely distinguish the interruptions that they have not learned recently. Interruption location procedures are of two sorts in particular; Misuse discovery and anomaly identification. Firewalls are utilized for interruption identification, yet they regularly fall flat in recognizing assaults that occur from inside of the association. An intrusion detection system (IDS) is a system that monitors network or overall system activities for malicious and unwanted activities or policy violations and produces reports to a management station. Intrusion detection system comes in variety of types and approaches for detecting suspicious traffic in different ways. There are network based (NIDS)[2] and host based (HIDS)[2] intrusion detection systems. NIDS is a network security system focusing on the attacks that come from the inside of the network. Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system. To overcome these disadvantages of firewalls, distinctive information mining procedures are utilized that handle interruptions happening from inside of the association.

Information mining systems have been effectively utilized for interruption discovery as a part of distinctive application zones like bioinformatics, securities exchange, web examination and so on. These systems extricate past obscure critical connections and examples from vast databases [3]. The extricated examples are then utilized as a premise to recognize new attacks. Information mining based IDS require less master learning, yet gives great execution and security. These frameworks are equipped for recognizing referred to and also obscure assaults from the systems. Diverse information mining methods like grouping, bunching and affiliation principle can be utilized for investigating the system activity and accordingly identifying interruption. Data mining is the way in which data is analyzed from different perspectives and it is summarized into useful information that can be used for gaining profit and cutting costs. Different data mining tools are present for analyzing data. Those tools allow users to analyze data from different dimensions and angles, classify, and summarize the relationships identified. Data mining is the method of finding correlations [8] or patterns [9] among many different fields in large datasets.

Attacks done by external malicious user are known as external attacks. Attacker is external to system. External attacker sends unwanted traffic to legitimate user. Due to traffic, legitimate user denies its expected services. In this case, user is said to be victim of DoS attack. Distributed Denial of Service (DDoS) [9] attack is extension to DoS attack. In DDoS attack, multiple slave computers are used to launch attack on legitimate user. Slave computers DDoS attack, launch attack by sending unwanted and malicious information to victim. Attacks done by internal and

authorized user are known as internal attacks. An insider attack in the system is a malicious and unwanted attack perpetrated on a network or computer system by a person with authorized user id and password. Authorized users that perform attacks have different malicious advantages over external attackers because they have authorized id and passwords to access the system and they are also familiar with network architecture and system policies. In addition, there is less security against insider attacks because in many organizations focus is on protection from external attacks. In remaining paper section II describes literature survey for system. Section III describes system architecture. Section IV includes results of system in tabular form. Section V shows performance graphs. Finally, section VI is dedicated to conclude paper.

II. LITERATURE SURVEY

Many researchers have worked on pattern matching algorithm for detecting SQL injection attacks. Most of the researchers concentrate on genetic algorithm for creating the rules with KDDCUP 99 dataset. The preparation and testing datasets gives the better consequences for all existing systems. Some import applications of soft computing techniques for Network Intrusion Detection is describe in this section. Several Genetic Algorithms (GAs) [11] and Genetic Programming (GP) [13] has been used for detecting intrusion detection of dissimilar kinds in different scenarios. Some uses GA for deriving classification rules. GAs used to choose required features and to determine the finest and negligible parameters of some core functions in which different AI method were used to derive gaining of rules. There are several papers related to IDS which has convinced level of impact in network security.

K. Marimuthu [2] has described solution for detection and prevention of SQL Injection Attack using Aho Corasick pattern matching algorithm. It is evaluated by using sample of attack patterns. Initial stage shows that the proposed solution is product of not false positive and false negative. The pattern matching process takes O (n) time. Bahare Pour [3] presents a classification of SQLIA based on vulnerability. Afterwards, divide the SQL injection and prevention methods to three different categories: static, dynamic and hybrid approaches. The paper explains different SQL detection and prevention techniques. We evaluated these techniques, with respect to deployment requirements. SQL injection payloads require a limited set of characters to fully exploit vulnerability. P. Murugeswari [4] has explained use of decision tree for detecting internal attacks. System call pattern is stored and log file is generated. Common system call patterns are filtered and similarity score is calculated. With the help of decisive rate threshold, it is identified whether the user is authorized account holder or not. Wei Li [5] describes the using genetic algorithm for intrusion detection method different detection technique. He has worked on TCP/IP layer. In the proposed arrangement is genetic algorithm rule bottom system which including crossover, mutation, fitness and assortment process and finally generate the rules for test data. The proposed algorithm captures the global semantic information using WordNet. WorldNet is a online lexical database for English language. The projected algorithm works in following manner. In this approach he discusses a methodology of applying genetic algorithm into network intrusion detection techniques. A overview of Intrusion Detection System (IDS), genetic algorithm are discussed. Factors affecting the GA are addressed in aspect. This accomplishment of genetic algorithm is unique as it considers equally temporal and spatial information of network associations during the encoding of the problem; therefore, it should be more helpful for identification of network anomalous behaviors. System used the assessment function in genetic algorithm. The evaluation function is one of the most imperative parameters in genetic algorithm. Finally system works on both KDD CUP 99 train as well as on training dataset with the suitable detection rate.

III. SYSTEM ARCHITECTURE

Proposed system includes both internal as well as external intrusion detection mechanisms. As shown in the Figure 1, system contains two algorithms to detect intrusion. System detects both internal as well as external attacks, in integrated way. In proposed system, when multiple computers are attached to each other, each system has its internal intrusion detection mechanism and for external intrusion detection FGA is used. In internal intrusion detection system, signature matching algorithm is used. When user fires query on database, it is checked using algorithm. In processing of algorithm, fired query is compared with stored signatures of malicious queries. If similarity index exceeds defined threshold then it is classified as anomalous query and notification is given to admin to take further actions.

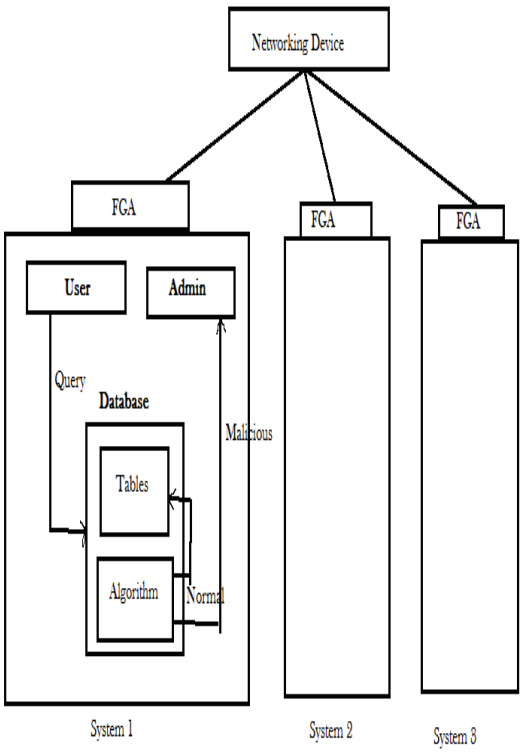


Figure 1. Intrusion detection system

A. Internal attack detection system

As classified above, user to root attacks are internal attacks to system. Internal attackers are the attackers which are authorized users of system but they perform malicious activities for some selfish purposes. Attackers can steal the id and password of authorized user and enter into the system for malicious purposes. Internal attacks include deleting some important files or data, alteration of some private files. SQL injection is one type of internal attack. In SQL injection, attacker enters into the system using authorized id and password and fires malicious queries on database. Database contains very important data about particular organization. By firing those malicious queries, data in database is altered or deleted.

For detecting these types of SQL injection attacks, algorithm 1 i.e signature matching algorithm is used. Some standard signatures of malicious queries are stored. When user fires query, it is compared with stored signatures. In static SQL injection detection, whole query is compared with stored signature if it matches with it then that query is classified as malicious otherwise normal. In dynamic SQL injection detection, similarity index of comparison is calculated. If similarity index exceeds defined threshold then it is classified as malicious query otherwise normal query. Threshold value is random value selected on basis of how much accurate result we want. If we define threshold value below 0.30 then more accurate results are obtained. Anomaly score is calculated by comparing fired query with stored signatures. Obtained value is anomaly score.

B. External attack detection system

Remote to local attacks are external attacks. In network, external attacker sends unwanted packets to legitimate machine in network. Due to traffic created by attacker, expected services of legitimate user are denied. This is denial of service attack in network. Attacker sends unwanted packets to legitimate user by forging source IP address of packet.

For detecting these types of external attacks, algorithm 2 i.e. Fuzzy Genetic Algorithm (FGA) is used. Using this algorithm, packets are classified into two groups i.e anomaly packets and normal packets. In FGA, rule creation is done by selecting population size and number of generations. Firstly random selection from population is done. After that cross mutation is done. Best probability attributes are selected and rules are generated. By comparing with created rules, packets are classified into normal or anomaly.

Algorithm 1: Pattern Matching Algorithm

Step1: Procedure SPMA(Query, SPL[ ])
Input: Query=User Generated Query
SPL[ ]=Static Pattern List with m AnomalyPattern
Step 2: For j = 1 to m do
Step 3: If (AC (Query, String.Length(Query), SPL[j][0]) = =0)then
Step 4: Calc anomaly score
Step 5: If ( ) Score Value Anomaly = Threshold
Step 6: then
Step 7: Retrun notification.
Step 8: Else
Step 9: Return Query- Accepted
Step 10: End If
Step 11: Else
Step 12: Return Query -Rejected
Step 13: End If
Step 14: End For
End Procedure

Algorithm 2: Fuzzy Genetic Algorithm.

Input: Test data and rule form rule pool , outsize fuzzy function with return a double value between 0 to 1.0
Output: Record with attack type

```
Step 1: for each record {
Step 2: for each rule{
Step 3: for each attribute{
Step 4: prob = fuzzy(Attribute);
Step 5: total prob = total prob + prob;}
Step 6: If (total prob > threshold){
Step 7: class is attack;
        true negative ++; }
        else{
        class is normal;
        true positive++;
            }
```

IV. RESULT TABLE

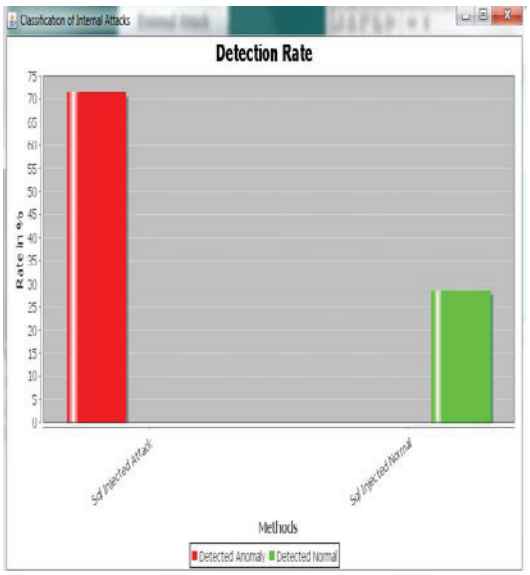
For detecting internal intrusion, when queries are fired they are classified as normal or anomaly. For external intrusion detection, when packets are received they are classified into normal and anomaly packets. Table 1 shows overall results of system

Table 1. Result table

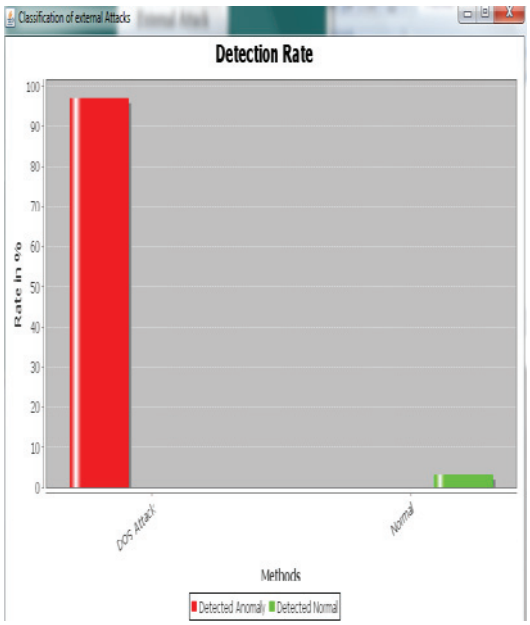
Algorithm	Total Records	Anomaly	Normal
Pattern Matching Algorithm	10	2	8
Fuzzy Genetic Algorithm	40	28	12

V. PERFORMANCE GRAPHS

a. Internal attack detection



b. External attack detection



VI. CONCLUSION AND FUTURE WORK

In this paper, different types of intrusion detection techniques are studied. Different techniques such as Neural Network, Clustering, Genetic algorithms are present to detect intrusion. Sometimes authorized users of system do intrusion for malicious purpose. The above mentioned attacks are hard to detect. Signature matching algorithm is used to detect internal intrusion. Fuzzy genetic algorithm is used to detect malicious packets in external attack detection. The proposed system helps to detect internal and external attackers in integrated way in a single system and block them. The proposed system provides better accuracy as compared with some of existing systems.

For further enhancement in system, we can focus to use of more than one algorithm to detect intrusion in single system.

REFERENCES

[1] P. Jongsuebsuk, N. Wattanapongsakorn, C. Charnsripinyo, “Real-Time Intrusion Detection with Fuzzy Genetic Algorithm ", in Proc. IEEE Int. Conf. Avail., Rel. Security, Vienna, Aus- tria, Apr. 2013

[2] Dr.M.Amutha Prabakar,M.KarthiKeyan, “An efficient tech nique for preventing SQL injection attack using pattern matching algorithm ", 2013 IEEE International Conference on Emerg- ing Trends in Computing, Communication and Nanotechnology (ICECCN 2013)

[3] Sayyed Mohammad, Sadegh Sajjadi, Bahare Tajalli Pour, “Study of SQL Injection Attacks and Countermeasures ", Inter- national Journal of Computer and Communication Engineering, Vol. 2, No. 5, September 2013

[4] S. Gajek, A. Sadeghi, C. Stuble, and M. Winandy, “Compartmented security for browsers Or how to thwart a phisher with trusted computing ", in Proc. IEEE Int. Conf. Avail., Rel. Security, Vienna, Austria, Apr. 2007, pp. 120127.

[5] C. Yue and H. Wang, “BogusBiter: A transparent protection against phishing attacks", ACM Trans. Int. Technol., vol. 10, no. 2, pp. 131, May 2010.

[6] Q. Chen, S. Abdelwahed, and A. Erradi, ”A model-based approach to self-protection in computing system ", in Proc. ACM Cloud Autonomic Comput. Conf., Miami, FL, USA, 2013, pp.110.

[7] F. Y. Leu, M. C. Li, J. C. Lin, and C. T. Yang, “Detection workload in a dynamic grid-based intrusion detection environment", J. Parallel Distrib. Comput., vol. 68, no. 4, pp. 427442, Apr. 2008.

[8] H. Lu, B. Zhao, X. Wang, and J. Su, “DiSig: Resource differentiation based malware behavioral concise signature generation", Inf. Commun. Technol.,vol. 7804, pp. 271284, 2013.

[9] Z. Shan, X. Wang, T. Chiueh, and X. Meng, “Safe side effects commitment for OS-level virtualization ", in Proc. ACM Int. Conf. Autonomic Comput., Karlsruhe, Germany, 2011, pp. 111120.

[10] Amit Kumar Pandey, “Securing web applications from applica tion level attacks", master thesis, 2007

[11] C.J. Ezeife, J. Dong, A.K. Aggarwal, “Sensor WebIDS: A Web Mining Intrusion Detection System", International Journal of Web Information Systems, volume 4, pp. 97-120, 2007

[12] S.Axelsson, “Intrusion detection systems: A survey and taxon- omy", Technical Report, Chalmers Univ., 2000

[13] Marhusin, M.F.; Cornforth, D.; Larkin, “An overview of recentadvances in intrusion detection", in proceeding of IEEE 8th International conference on computer and information technology CIT, 2008

[14] S. F. Yusufovna., “Integrating Intrusion Detection System andData Mining ", International Symposium on Ubiquitous Multi- media Computing, 2008

[15] J.Gmez and E. Len, “A fuzzy set/rule distance for evolving fuzzyanomaly detectors", IEEE International Conference on FuzzySystems, ART. No. 1682017, pp. 2286-2292.