A Pattern Matching Algorithm for Reducing False Positive in Signature Based Intrusion Detection System

[1]T. Sree Kala, [2]Dr. A. Christy

[1] Research Scholar, Bharathiar University, Coimbatore, India
[2]Professor & Principal, St.Mary's School of Management Studies, Chennai, India
[1]sreekalatm@gmail.com
[2]ac.christy@gmail.com

***Abstract*-Nowadays the organizations are facing the number of threats every day in the form of viruses and attack etc. Since many different mechanisms were preferred by organizations in the form of intrusion detection and prevention system to protect its organizations from these kinds of attacks. Intrusion Detection System (IDS) is considered as a system integrated with intelligent subsystems. In this paper the signature based intrusion detection system is discussed. There are different pattern matching algorithms available to detect intrusion. Brute force and Knuth-Morris-Pratt are the single keyword pattern matching algorithms. If one or more occurrence of pattern present in the input text, then there is an intrusion and the intrusion alarm will be sent. The occurrence of false alarm will be high in intrusion detection. In this paper the string matching algorithm to reduce the percentage of false alarm will be discussed.**

Keyword - Intrusion Detection System, subsystem, Threats, Viruses, attacks, signature, anomaly, false alarm

I. INTRODUCTION

Intrusion Detection was developed in the late 1990s to identify and report the attack, it detected hostile traffic and sent alerts but it failed to stop the attacks [1]. It detects the hacker’s attacks and network worms began to affect the internet. Intrusion Detection is passive, that is not able to detect all malicious activities most of the time and it has control restriction to stop traffic inbound - outbound from

attacking. It was only capable to detect attack actions, without prevention action.

According to Anderson [2], an intrusion attempt or a threat is a unauthorized attempt to (i) accessing information, (ii) using information, (iii) render a system unreliable or unusable. For example, (a) *Denial of Service (DoS)* attack attempts to block access to system or network resources, which are needed to works correctly during processing; (b) *Worms and viruses* A self-replicating program exploit the system without any knowledge or permission from the users through the network; and (c*) Network Attack* Illegally using user accounts and obtain privileged access to a host by taking advantages of known vulnerabilities.

Anderson [2] categorizes intruders into two types: external and internal. External intruders are unauthorized users of the machines they attack. Internal intruders have permission to access the system, but do not have rights for accessing the superuser mode. There are various classes of intrusions or attacks [4], [5].

II. INTRUSION DETECTION SYSTEM

Intrusion is a set of actions used to provide the security of computer and network components in terms of confidentiality, integrity and availability [3]. This can be completed by an inside or outside agent to gain unauthorized entry and control of the security mechanism. Intrusion detection systems (IDSs) provide well-established mechanisms, which collect and evaluate data from various areas within a system or a network to identify possible security breaches. An IDS use openness valuation to assess the security of a host or a network. Intrusion detection works on the hypothesis that intrusion activities are deviates from normal system activities and thus detectable. The structure of Intrusion Detection System is shown in Fig.1.

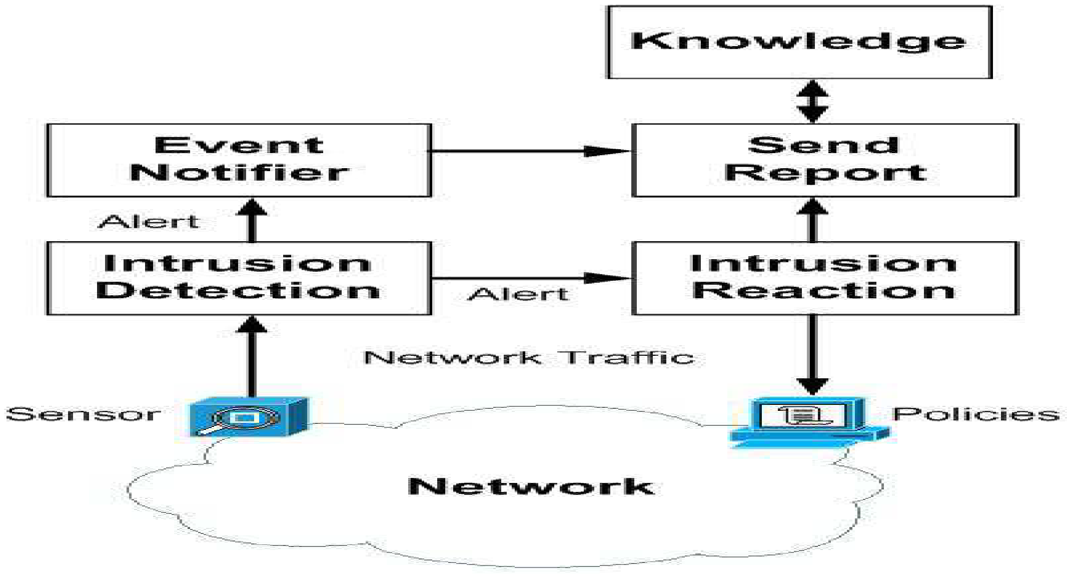

Fig.1: Intrusion Detection System

Intrusion detection provides the following:

a. Monitoring and analysis of user, system and network activity.

b. Checking and generating reports of vulnerabilities.

c. File integrity and availability of critical data files.

d. Recognizing patterns based on the matching to known attacks using statistical analysis of activity.

e. Check abnormal behavior activity.

f. Analyzing Operating systems and compare the stable state.

III. CLASSIFICATION OF IDS

IDS can be properly categorized into two specific patterns as necessary. We can discuss the two types as following.The first classification is based on the IDS to be placed in the organization. It will explain how the IDS can be deployed in the real rime applications. It can be classified into three groups.

i. Host Based Intrusion Detection System (HIDS)

ii.Network Based Intrusion Detection System. (NIDS)

iii. Hybrid Based Intrusion Detection System

A. Host Based Intrusion Detection System:

A HIDS observes and analyzes the internal parts of a computing system rather than its external peripheral interfaces [12]. It works combined with a software agent on a host. The IDS sensors are located at choke points and used to monitor the host. The most modern HIDS are proposed as host based applications running in the background of critical, sensitive hosts. The examples are Mail Servers, DNS Servers, Web servers, Database servers, etc. It detects the interventions by analyze the system calls, log files and file system modifications. The good example for open source application-based IDS are OSSEC [13]. The operating system enforces the HIDS as an agent that monitors whether anything internal or external resources have circumvented the security policy. The Fig.2 explains the typical HIDS.

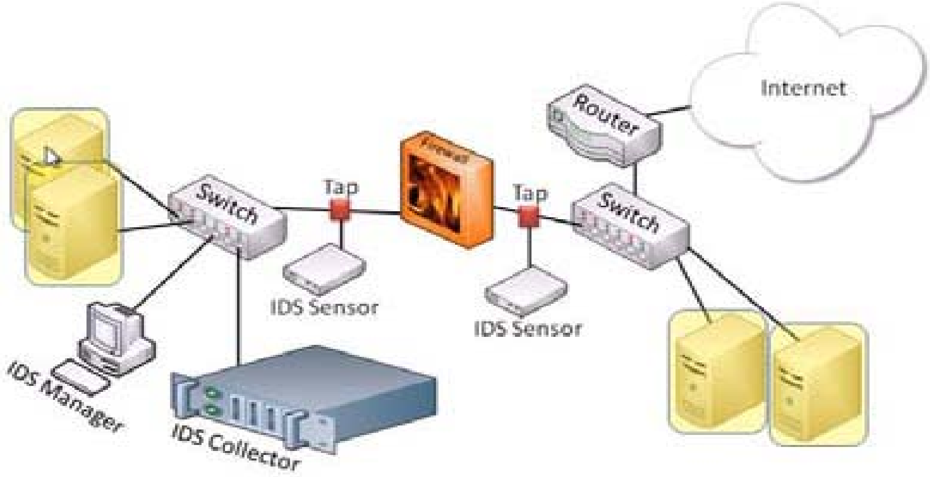

Fig.2: Host based Intrusion detection system

B. Network Intrusion Detection Systems:

A NIDS is an independent platform that deals with detecting intrusions in network data and multiple hosts. The NIDS used the technique to identify the intrusions from the network traffic using the network hub, network switch or network tap. The IDS sensors are used to monitor the network to capture the individual packets in the anomalous patterns in the demilitarized zone. The anomalies are launched by the attackers from outside, who want to be access the network unauthorized to steal the data and interrupt the network traffic. The NIDS used the technique of 'port scan' [14] and various tools to scan the incoming packets and identify anomaly subsequences [12]. After detecting the intrusions the NIDS provides the details about them that are in packet level or in outgoing network. A good software example for a NIDS is SNORT. The Fig.3 explains the typical NIDS.

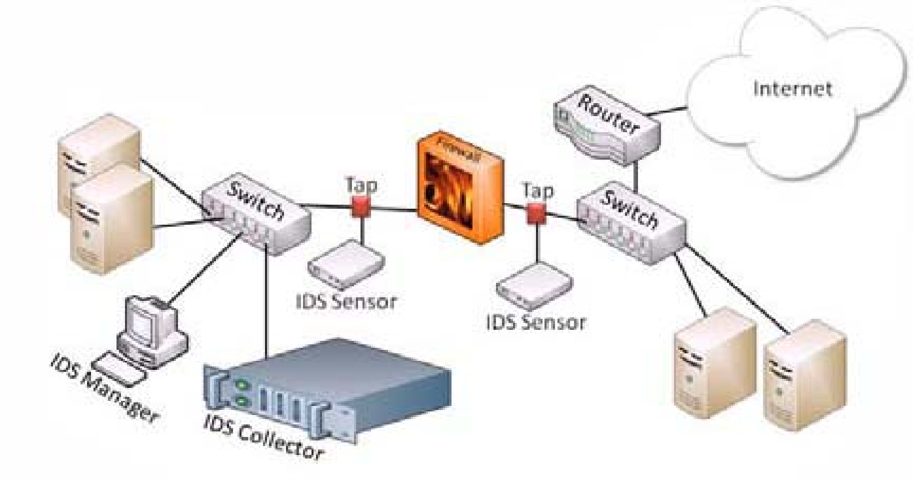

Fig.3: Network Intrusion Detection System

C. Hybrid Intrusion Detection Systems:

Both types of Intrusion Detection Systems can also be combined, and that systems are called Hybrid Intrusion Detection Systems. The system achieves the benefits of both misuse and anomaly-based detection techniques. It is used to detect known as well as unknown attacks.

IV. METHODOLOGIES OF INTRUSION DETECTION

There are two approaches available based on analysis of the technique used. Therefore the intrusion detection system classified based on analysis pattern. This scheme pattern can be generally divided into two units:

a. Mis-use Based Intrusion Detection System

b. Anomaly Based Intrusion Detection System.

A. *Mis-Use Based Intrusion Detection System:*

The detection of intrusion is based on a set of rules or looked for known attack signatures by its reference. This type of identification is known as knowledge-based or misuse detection IDS. The signatures are screened by IDS by packet by packet in the network and compares with preconfigured and predetermined attack patterns. When a new attack is encountered the experts or programs have to identify usual patterns and it can be made into signature. The ways to write a signature to incorporate all possible deviations of the pertinent attack is a challenging task. The examples of signatures are

- Failed to login on the specific host.
- Buffer overflow of IP packet.
- SYN flood Does attack.

The different types included in the signature based intrusion detection system are:

- Expert system
- Signature Analysis
- Petri Nets
- State Transition

B. *Anomaly-Based Intrusion Detection System*:

All interfering activities are known as anomalous. Anomaly-based IDSs checks whether the system violates from the typical behavior profile, and detected by the statistical analysis. Some anomalous activities are not intrusion but may be identified as intrusive. These types of anomalies are known as false positive. The examples for anomalous activities are

- Masquerade attacks in security control system
- Denial of service attacks and Leakage
- Malicious access and
- Interruptions in security constraints
- Using special privileges[15]

The different types included in the Anomaly intrusion detection system are:

- Statistical Based IDS
- Expert system IDS.
- Neural Networks IDS.
- User Intention Identification System.
- Computer immunology IDS.
- Data Mining based IDS.

V. ALGORITHMS FOR SIGNATURE BASED INTUSION DETECTION

Single keyword pattern matching algorithms are detecting the payload intrusion. The pattern is matched with the input text. The pattern and input are fixed and finite non empty alphabet. The algorithm produces the result of all occurrences of the pattern in the input text.

A. *Brute force algorithm*:

Brute force algorithm is a very well-known string matching algorithm. It checks the position of text from 0 to m-n with a pattern of size m. The procedure is comparing each character in pattern with the corresponding character in the text. If there is any unmatched character, then the data is intruded [17].

Example:

Input : UPDATEVIRUSSIGNATURE

Pattern:VIRUS

1) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisions made

2) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisions made

3) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisions made

4) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisons made

5) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisons made

6) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisons made

7) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisons made

Pattern is found. The input is corrupted by intruders. 35 comparisons needed to find the interruption.

B. Knuth-Morris-Pratt algorithm:

Knuth proposed a string matching algorithm that search string into a finite state machine. It avoids the comparisons with the elements of _S_ , which is involved in comparisons with some elements of the pattern _p_ to be matched. A matching time of O(n) is achieved by avoiding the comparisons with the elements that have previously been involved in comparison with some element of pattern.

Example:

Input : UPDATEVIRUSSIGNATURE

Pattern:VIRUS

1) UPDATEVIRUSSIGNATURE

VIRUS 5 comparisions made

2) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisions made

3) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisions made

4) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisons made

5) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisons made

6) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisons made

7) UPDATEVIRUSSIGNATURE

VIRUS 1 comparisons made

Pattern is found after 11 comparisions.

C. Less False Alarm Algorithm:

False positive is the alarm that triggered, when the attack is not being happened. The good IDS should identify the real attacks and it must have less number of false positive. Generally 96%of alarms are false positive. The LFA algorithm [16] will reduce the count of false alarm. In pattern matching algorithm the pattern is defined as intrusion. But sometimes it may be data. The IDS generate alarm each and every time matching the pattern. If the pattern occurs more than two times in the input, the possibility for the intrusion is higher. In this algorithm, the alarm will be generated if the pattern is repeated more than twice. This will create a table to reduce the number of comparisons.

Example 1:

Input: UPDATE VIRUS SIGNATURE

Pattern:VIRUS

1) **U**PDATE VIRUS SIGNATURE

VIRUS 1 comparison

2) U**P**DATE VIRUS SIGNATURE

VIRUS 1 comparison

3) UP**D**ATE VIRUS SIGNATURE

VIRUS 1 comparison

4) UPD**A**TE VIRUS SIGNATURE

VIRUS 1 comparison

5) UPDA**T**E VIRUS SIGNATURE

VIRUS 1 comparison

6) UPDAT**E** VIRUS SIGNATURE

VIRUS 1 comparison

7) UPDATE **V**IRUS SIGNATURE

VIRUS 5 comparisons count +1

8) UPDATE VIRUS SIGNATURE

VIRUS 9 comparisons

Finally count=1, Which is<=3 , then the data is not intruded, and Print “normal data”.

Example 2

String: VIRUS SIGN VIRUS SIGN VIRUS SIGN

Pattern: VIRUS

1) VIRUS SIGN VIRUS SIGN VIRUS SIGN

VIRUS 5 comparisons count+1,n-m

2) VIRUS SIGN VIRUS SIGN VIRUS SIGN

VIRUS 1 comparisons

3) VIRUS SIGN VIRUS SIGN VIRUS SIGN

VIRUS 5 comparisons count+1,n-m

4) VIRUS SIGN VIRUS SIGN VIRUS SIGN

VIRUS 1 comparisons

5) VIRUS SIGN VIRUS SIGN VIRUS SIGN

VIRUS 5 comparisons count+1,n-m

6) UPDAT**E** VIRUS SIGNATURE

VIRUS 1 comparisons

Finally count=3,

Which is >=3, then the data is corrupted, and Print “Intrusion”.

VI. CONCLUSION

In this paper, we have discussed about different types of intrusion detection against the malicious acts and attacks to secure an organization from threat. Hence we have given a classification scheme for these intrusion detection and prevention. Network based Intrusion detection system can detect small attacks. Signature based IDS play an important role in NIDS. The IDS cannot able to detect the new attacks if the IDS was not updated the signature of new pattern periodically. The main challenge of IDS is reduce false alarm. LFA algorithm will help to suppress the false alarm. It will generate the alarm only, when the pattern present more than two times in input. Otherwise ignore it. In future the improved solution for the issues can be discussed.

REFERENCES

[1] Y. Weinsberg. S Tzur-David, D. Doley, and T.Anker, “High Performance String Matching Algorithm for a Network Intrusion Prevention System(NIPS),” High Performance Switching and Routing, IEEE, 2006, pp. 147-153.

[2] J. P. Anderson, “Computer Security Threat Monitoring and Surveillance,” Jammes P Anderson and Co, Fort ashington, Pennsylvania, Tech Rep, April 1980.

[3] R. Heady, G. Luger, A. Maccabe, and M. Servilla, “The Architecture of a Network Level Intrusion Detection System,” Computer Science Department, University of New Mexico, Tech. Rep. TR-90, 1990.

[4] H. G. Kayacik, A.N. Zincir-Heywood, and M.I. Heywood, “Selecting Features for Intrusion Detection: A Feature Relevance Analysis on KDD 99 Intrusion Detection Datasets,” in Proc. 3rd Annual conference on Privacy, Security and Trust, October 2005.

[5] 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.

[6] A. Singhal, Data Warehousing and Data Mining Techniques for Cyber Security, Advances in Information Security Springer, 2007.

[7] D. Stiawan, A.H. Abdullah, and M.Y. Idris, “The Trends of Intrusion Prevention System Network,” International Conference Education Technology and Computer, Shanghai, China: IEEE, 2010,pp. 217-221.

[8] S.H. Oh and W.K. Lee,” An anomaly intrusion detection method by clustering normal user Behavior,” Computers & Security, vol. 22, 2003, pp. 596-612.

[9] A.D. Todd, R.A. Raines, R.O. Baldwin, B.E. Mullins, and S.K. Rogers, “Alert Verification Evaluation Through Server Response Forging,” INCS, vol. 4637/2007, 2007, pp.256-275.

[10] H. Artail, H. Safa, M. Sraj, I. Kuwatly, and Z. Almasri, “A hybrid honeypot Framework for improving intrusion detection systems in protecting organizational networks,” Computer & Security, vol. 25, 2006 pp. 274-288.

[11] P. Garcia-Teodora, J. Dian-Verdejo, G. Macia- Femandez, and E. Vazquez, “Anomaly- based network intrusion detection Techniques, systems and challenges,” Computer & Security, vol. 28, 2009, pp. 18-28.

[12] Wikipedia, “Intrusion Detection System,” http://en.wikipedia.org/wiki/Inrusion-detection system, Feb 2009.

[13] OSSEC(Observing System Science Executive Council) OSS.Homepage of OSSEC, 2011, http://www.ossec.net/.online;

[14] H Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, “Surveying Port Scans and Their Detection Methodologies,” The Computer Journal, vol. 54, no. 10, pp. 1565-1581, October 2011.

[15] Joseph Migga Kizza. Computer Network Security, Springer, 2005, Part III, 315-346. DOI: 10.1007/0-387-25228-2 12.

[16] Lata, ashyap Indu, “Novel Algorithm for Intrusion Detection System,” IJARCCE, vol 2, pp. 2104-2110, May 2013.

[17] Siddharth Saha,”Network Intrusion Detection System Using String Matching,” Internet,1-46,2010.