ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 4, April 2013

Post-Attack Detection Using Log Files Analysis

Bharat Sampatrao Borkar[1], Apurva Suresh Patil[2]

Assistant Professor, Dept. of IT, AVCOE, Sangamner, Maharashtra, India[1]

Student of M. E. (IT), AVCOE Sangamner, Maharashtra., India[2]

Abstract: Security has become a most important issue in recent years, for that much intrusion detection systems have been proposed. Though there are lots of systems available we still need a system which will timely detect the intrusions. Proposed system is a host based intrusion detection system can be called is Post-Attack intrusion detection. We are investigating the system log files which contain the log of all system calls. The system has two main features.
1) It reduces the time to locate a particular log with intruder activities by factoring it.
2) A classifier which will classify the normal behaviour form malicious one.
To factor the log files sequitur method is used which will reduce the size of log, and a classifier is the main part of system which is using a HMM (Hidden Markov Model) and k-means to classify normal and abnormal behaviour.

Keywords: Interface Anomaly Detection, Host-based Intrusion detection, HMM (Hidden Markov Model), Sequitur.

I. INTRODUCTION

An Intrusion detection system is one of the possible solutions to timely detect the intrusion and alert the user. Intrusion detection has been studied for approximately 20 years. Intrusion is nothing but any activity that tries to violate the integrity, confidentiality by gaining unauthorized access to your system .to avoid such an unauthorized access many mechanisms are introduced such as passwords, encryption. But again these are prevention mechanism none of them is capable to detect the intrusion, i.e. is anybody having your password can access the system and it fails to detect that it is an intruder and not a legitimate user. So to overcome this drawback the intrusion detection systems are introduced. So the detection is based on the beliefs that an intruder's behaviour will be noticeably different from that of a legitimate user and that many unauthorized actions will be detectable .The IDS are broadly classified as host based IDS(Intrusion Detection System)and network IDS(NIDS). HIDS is limited to detect intruder on single host. It observes the underlying operating system as it executes system call. NIDS is used to detect the intrusion over network.

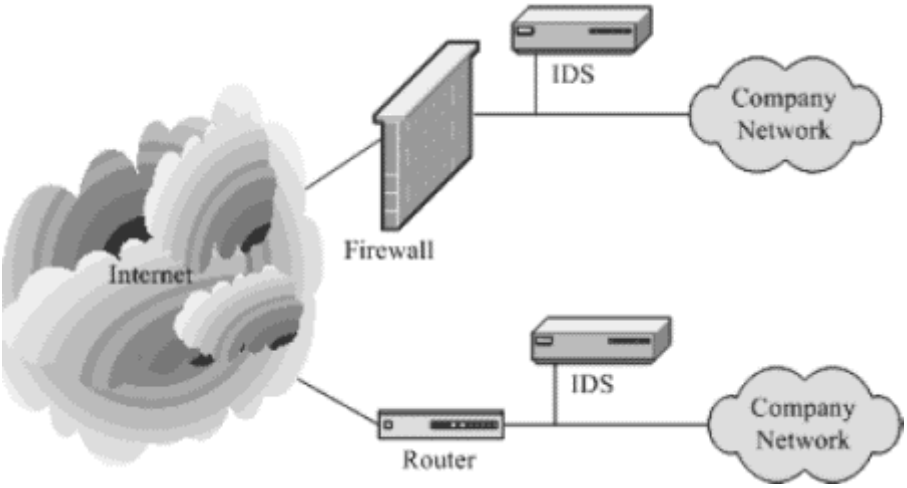

Fig 1: Position of IDS

Above fig 1 denotes the position of IDS in network. A network based intrusion detection system captures packets passing through the network, and later the captured information is analysed and detection is performed. Alternately IDS can be classified into two methodologies: Anomaly detection and misuse (signature-based) detection. In Anomaly detection the profile of user behaviour is created and, if some difference occurs in style of execution of user i.e. consuming more resources, more bandwidth, key stroking speed, then this deviation from normal behaviour is observed and the user or system is considered as intruder. On the other hand in misuse detection signatures are created these signature can be called rules. the signature is nothing but pattern of attack e.g. Consider an Network Intrusion detection system which detects intruder in network for that a signature can be having attribute as source Ip of packet ,Destination IP address, source port number, destination port number and so on. Now these type of signatures are saved in database

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 4, April 2013

or say in a rule file and whenever any intruder try to exploit, its signature is matched with the one from database. If the signature matches the system gives result as intruder. "Snort" is type of Rule based IDS.

Many times it happens that intruder can bypass the IDS and harm the system. E.g. in case of zero day attack. Therefore it becomes essential to Analyse the system or host., From this analysis we can get information like how the intruder gained access to this particular system as well as what actions did he perform The term post-mortem intrusion detection is somewhat similar to intrusion detection. The only thing that is different is the ids alarms user whenever intruder found, but this system monitors offline system information and then detects the exploit. So, the problem is using a log file find out the exact information about the exploit if any.

II. RELATED WORK

To address the problem we need to understand the advantage and disadvantage of existing intrusion detection systems. As explained in Introduction Section the intrusion detection is mainly classified as Anomaly detection and misuse detection. On the bases of their detection technique And alternately they can be classified as host-based intrusion detection and Network based Intrusion detection system. Here our main concern is anomaly based IDS, so for that we should do some study on anomaly based intrusion detection systems. The research in this area is three approaches as

1. System-based detection
2. Specification –based detection
3. learning-based detection

A. System-Based Detection:

The System-Based Intrusion detection we can say that is completely depends on Audit records. As we all know the audit record contains all the information like CPU/memory usage, login attempts (valid invalid), amount of I/O data, and so on. In System-based detection the events form audit trail are observed and the normal behaviour pattern .in early 1980's these systems were into use .

B. Specification based

The application specific policies are the base of specification based intrusion detection. It may depend on the software to generate the policies or can be specified manually [18] the specifications describe the intended behaviour of program. Monitoring programs involves detection of deviation from these specifications the main advantage is the attacks can be detected even if it has not occurred previously.

C. Learning Based

It has two major areas for research as:

- Detection using Sequences of System Calls: According to sequence of system calls, any unusual system calls are observed in this technique. One of the approach is

  - Immune system based detection:

In [2] they have presented time-delay embedding (called tide, for short).this is based on the principles of a natural immune system. IT is an early mechanism for intrusion detection. The extension to tide is presented later [3], which they named as sequence tide(called stide).In stide a database of contiguous sequence of system call is maintained. Then, after giving training set the stide first divide it into fixed length sequences and then compare these with database.

Another variation of stide is introduced as T-stide it is just as stide but it considers the frequency of appearance of each system call. Second approach for detection using system call is:

- HMM-based approaches:

This is another anomaly based approach for intrusion detection using HMMs[5][6]. In this the number of states in the hmm was fixed which corresponds to number of system calls Baum–Welch algorithm was applied to compute the transition and symbol probabilities. a training session is conducted in which the trace is read,. only one system call is considered at a time, and then decision is made using HMM , the number of mismatches is used to determine whether the session is intrusion or not

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 4, April 2013

- Detection through the Analysis of System Calls' Arguments: Many IDS failed to detect mimicry attack; Kruger et al. [4] pointed out that to avoid that one should take full advantage of the information in system call traces. New system has proposed that combines many models that learn the characteristics of real system call arguments. Every model is employed, first, to derive a profile for a given style of argument, and, then, to yield an anomaly metric for a given system call. The foremost drawback of the mechanism given in [4] is that it is not scalable.

Later on an intrusion detector based on the analysis of system call arguments is proposed. The technique here attempts to capture temporal properties about the argument values of different system calls. The mechanism is extended which had few differences. first ,it makes use of models on clusters, each of which groups together invocations of the system call with similar arguments. Second, it adds program flow by modeling program execution using a Markov chain on system call clusters. and few more modifications improved the system very well. This is all about the research on Anomaly Based detection techniques. In next section we outline the concept of Post-Attack IDS.

III. PROPOSED SYSTEM

Our approach of Post-Attack intrusion detection is an Anomaly, learning-based, host-based intrusion detection model. The theory behind it is, the attack takes the form of an abnormal sequence of system calls. In our assumption the attacker has bypassed the online intrusion detection system, if any and there is a set of logs, where activities of the Intrusion have been recorded, in the form of a trace of system calls. We aim to develop a method that is capable of pinpointing in one such log where the execution of an exploit should exist. Using which an intruder has upgraded his privileges to those of a, for example, system administrator. Building a successful Post-Attack intrusion detection system we must address following key issues. First, accuracy: the correctness of the profile directly affects the performance of anomaly detection system, which is used to capture ordinary system behavior. Building such a profile is difficult in general, because normal and abnormal system behaviors have very less difference in between them, and hence, it is difficult to put them apart. Second, scalability: computer activity very quickly consumes gigabytes of information [5].

Two important observations are First, although huge, log files contain a number of repetitive, spurious information, which goes to Post-Attack intrusion detection Therefore, we factor out repetition, while attempting to identify relevant information although there is possibility of losing key information Second, although Post-Attack intrusion detection involves analyzing a complete log file, the construction of a good profile for ordinary behavior might be achieved, even though we restrict ourselves to only some parts of it.

Architecture for this proposed system consists of two most important things i.e. first is constructing a model of normal behavior, and second is the mechanism for Post-Attack intrusion detection.

IV. ARCHITECTURE

Fig.2 shows our approach for the construction of the model for ordinary behavior and fig 3 shows the mechanism for Post-Attack intrusion detection. Both tasks are approached using a similar procedure. To build a model for ordinary behavior, at first, we factor out repetitive Behavior in a collection of attack-free (ordinary) log files. As a result, a compressed version of all log files is obtained, and, seconds, a relation of the sequences of most common occurrence across all these logs, which we call repetitive sequences. In the next step of detection we follow a 100-size and 100-step sliding-window approach to analyze every reduced log: starting at the first position of the log, we retrieve a window of size 100, then characterize each window using an attribute vector, and then slide the window a step.

For post-attack intrusion detection, we have considered that across repetitive sequences can be used to build a compressor on an input test log file. Thus, in the first step, we reduce the size of test log by simply applying the across repetitive sequences, as shown in the reduction model in Fig. 3. Then, the output, reduced file is subject to a sliding window process, and, lastly, we apply the detection model; this step actually consists of comparing the average model built from ordinary activity, as shown in the detection model in Fig. 3, with the one built from the current testing session. If they do not differ significantly, the associated window is marked normal; otherwise, it is marked abnormal.

ISSN: 2319-8753
International Journal of Innovative Research in Science, Engineering and Technology
Vol. 2, Issue 4, April 2013

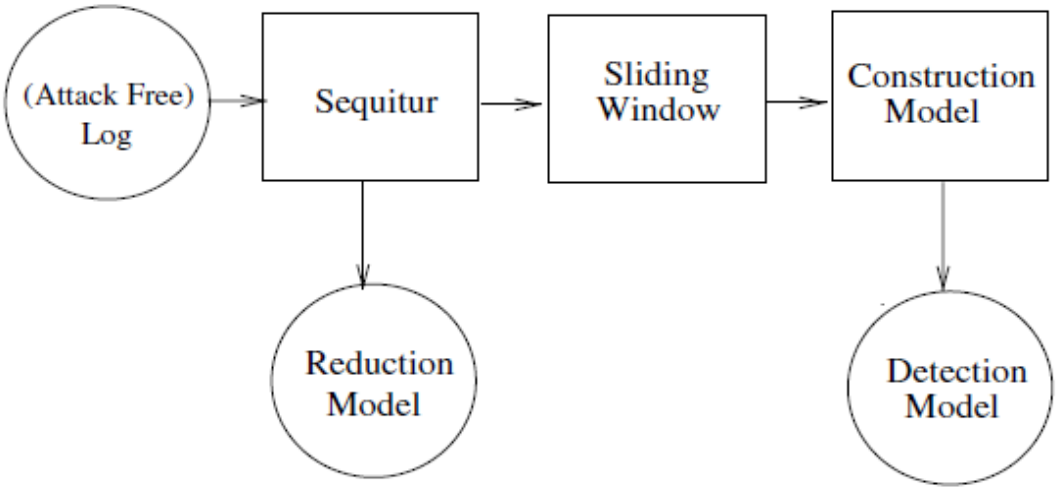

Fig 2: Model for construction of normal behaviour

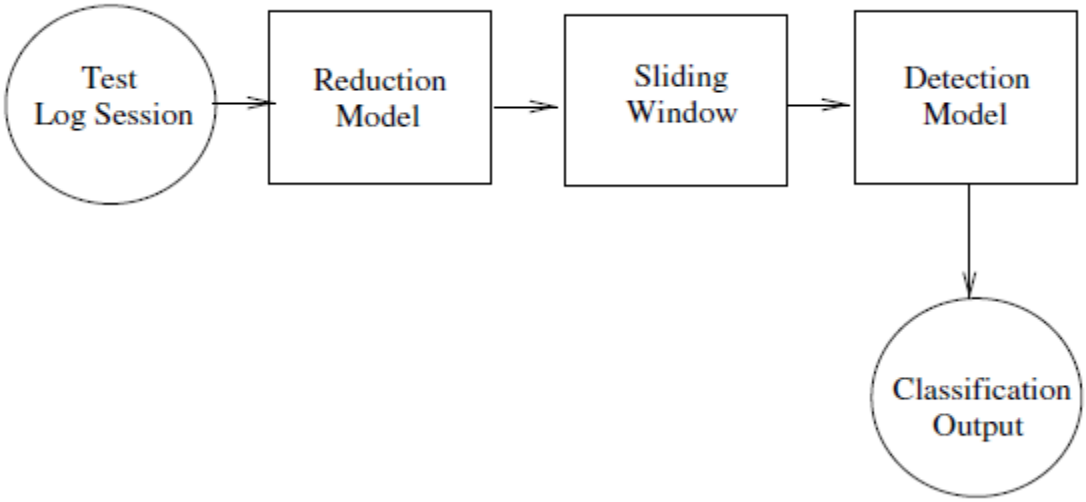

Fig 3:Post-Attack intrusion detection mechanism

For post-attack intrusion detection, we have considered that across repetitive sequences can be used to build a compressor on an input test log file. Thus, in the first step, we reduce the size of test log by simply applying the across repetitive sequences, as shown in the reduction model in Fig. 3. Then, the output, reduced file is subject to a sliding window process, and, lastly, we apply the detection model; this step actually consists of comparing the average model built from ordinary activity, as shown in the detection model in Fig. 3, with the one built from the current testing session. If they do not differ significantly, the associated window is marked normal; otherwise, it is marked abnormal.

V. CONCLUSION

In this paper, we have presented the post-attack intrusion detection system, which is a host based and anomaly based intrusion detection. It consists on pinpointing the execution of an exploit, if any Post-attack intrusion detection is important for computer post-mortems, because it speeds up the process of gathering evidence of an intrusion. Secondly it also speeds up the process of building an attack signature. An attack signature is very valuable for IDS construction, especially in the context of commercial IDS. Post-attack intrusion detection is very complex. As a future scope of this system we can build the same system for Network Intrusion Detection.

REFERENCES

[1] M. C. Ko, M. Ruschitzka, and K.-N. Levitt, *"Execution monitoring of security-critical programs in distributed systems: A specification-based approach,"* in Proc. IEEE Symp. Security Privacy, May 1997, pp. 175– 187.
[2] Z. Li and A. Das, *"Analyzing and evaluating dynamics in stide performance for intrusion detection,"* J. Knowl.-Based Syst., vol. 19, no. 7, pp. 576–591, 2006.
[3] S. A. Hofmeyr, S. Forrest, and A. Somayaji, *"Intrusion detection using sequences of system calls,"* J. Comput. Security, vol. 6, no. 3, pp. 151–180, 1998.
[4] C. Kr¨ugel, D. Mutz, F. Valeur, and G. Vigna, *"On the detection of anomalous system call arguments,"* in Proc. 8th Eur. Symp. Res. Comput. Security, 2003, vol. 2808, pp. 326–343.
[5] C. Warrender, S. Forrest, and B. A. Pearlmutter, *"Detecting intrusions using system calls:Alternative data models*," in Proc. IEEE Symp. Security Privacy, 1999, pp. 133–145.
[6] D. Yeung and Y. Ding, *"Host-based intrusion detection using dynamic and static behavioral models,"* Pattern Recognit., vol. 36, no. 1, pp. 229–243,2003.

BIOGRAPHY

Mr, Bharat Sampatrao Borkar received his B.E and M.E degrees in Computer Science and Engineering. He is a member ISTE, ACM, IE. He is working as Assistant Professor at AVCOE, Sangamner. His research interests includes Image processing and Intrusion detection.

Ms, Apurva Suresh Patil completed her B.E. in information Technology from Pune University, in 2012. She is currently doing his M.E. in Information Technology from University of Pune.