Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 09, September-2020

Empowering Intrusion Detection System in Cloud based Applications

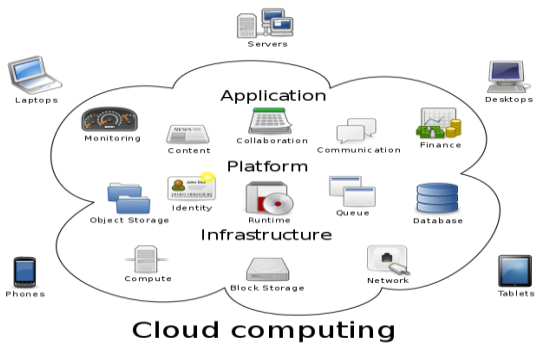
Mr. Ranjith S
Assistant Professor ,
SCMS School of Technology & Management

***Abstract*:-Today, Most of the IT organization prefers Cloud Computing as their choice because of its flexibility and pay-per-use based services to its users. However, the security and privacy is a major challenge in its success because of its distributed and open architecture which is vulnerable to intruders. Intrusion Detection system is the most commonly used mechanism to detect attacks on cloud based applications. IDS monitor the applications hosted in the cloud and alerts when an attack attempt is detected. This paper provides the method that empowers the Intrusion Detection System to detect attacks in cloud based applications and classify through clustering algorithm. This approach can create normality models of isolated user sessions that include both the web front-end (HTTP) and back-end (File or SQL) network transactions with respect to Data volumes and Classify them. The project implements a lightweight virtualization technique to assign each user's web session to a dedicated container, an isolated virtual computing environment. The system then use clustering algorithm to accurately associate each web request with the subsequent DB queries and build a causal mapping profile by taking both the cloud server and DB traffic into account. Finally the system is able to detect attacks both in static and dynamic websites with 100 percent accuracy by reducing the processing time.**

Keywords : Intrusion detection system , Cloud based applications, K-means clustering, double guard, Virtualization techniques

I. INTRODUCTION

Cloud computing is an emerging technology adopted by organizations of all scale due to its low-cost and pay-as-you-go structure. It has revolutionized the IT Industry with its unique and ubiquitous features. Organization prefers cloud as it replaces the high price infrastructure and need of maintenance. It provides a variety of services to end users in on-demand basis. 1. Software as a Service, 2.Platform as a service 3.Infrastructure as a service.



Intrusion detection techniques are used in any computing environment as a layer of defense. The basic aim is to detect any malicious activity well before any significant harm is possible. The general idea is to detect and identify attacks by either analyzing system artifacts (such as log files, process lists, etc.), or by keeping track of network traffic. Two main approaches used are signature based detection and anomaly-based detection. Signature based detection works by defining patterns of known attack signatures. If the system is found to be processing any code similar to those signatures, it is detected suspicious and marked as an intrusion. On the other hand, anomaly based detection works by analyzing activities performed on the system. Initially, a profile for a particular system is created by recording normal activities (e.g., by setting thresholds for normal bandwidth usage). If later on, the system's behavior is analyzed as anomalous to the profile defined, it is marked as an intrusion. Whereas signature-based detection techniques (also called misuse pattern matching) cannot detect unknown attacks, anomaly based techniques usually result in huge false positives or negatives.

II RELATED WORKS

Intrusion detection system in cloud computing : Challenges and Opportunities , This paper provides an overview of different intrusions in cloud. Then, analyzing some existing cloud based intrusion detection systems (IDS) with respect to their type, positioning, detection time, detection technique, data source and attacks they can detect. The analysis also provides limitations of each technique to evaluate whether they fulfill the security requirements of cloud computing environment or not. We emphasize the deployment of IDS that uses multiple detection methods to cope with security challenges in cloud.

Anomaly-Based Intrusion Detection System [7], is a system for detecting computer intrusions and misuse by monitoring system activity and classifying it as either *normal* or *anomalous*. The classification is based on heuristics or rules, rather than patterns or signatures, and will detect any type of misuse that falls out of normal system operation. This is as opposed to signature based systems which can only detect attacks for which a signature has previously been created.

Intrusion Recovery for Database-backed Web Applications [8] In this paper Users or administrators must manually inspect the application for signs of an attack that exploited the vulnerability, and if an attack is found, they must track down the attacker's actions

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 09, September-2020

and repair the damage by hand. When an administrator learns of security vulnerability in a web application, he or she can use WARP to check whether that vulnerability was recently exploited, and to recover from any resulting intrusions.

III INTRUSIONS IN CLOUD

A Attacks on hypervisor

An attacker may successfully control the virtual machines by compromising the hypervisor. The most common attacks on virtual layer are SubVir [8], BLUEPILL [9], and DKSM [10] which enable hackers to supervise host through hypervisor. Attackers target the hypervisor or VMs to access them by exploiting the zero-day vulnerabilities in virtual machines [11], prior to the developers' awareness about such exploits [3]. The exploitation of a zero-day vulnerability in the HyperVM application caused damage to several websites based on virtual server

B U2R attacks(User to Root)

The attacker uses password sniffing to access a genuine user's account which enables him to obtain root privileges to a system by exploiting vulnerabilities, e.g. Root shells can be created by using Buffer overflows from a root-level process. In the cloud scenario, attacker achieves root privileges of host or VMs by first getting access to legal user instances. This attack violates the integrity of cloud based systems

C. Insider Attack

The attackers are the authorized users who try to obtain and misuse the privileges that are either assigned or not assigned to them officially. This attack is closely related to trust since insiders may reveal secrets to opponents, e.g. Amazon Elastic Compute Cloud (EC2) suffered from an internal DoS attack. This attack breaches the confidentiality of cloud users

D. Backdoor channel attacks

Hackers can remotely access the infected machines by exploiting this passive attack to compromise the confidentiality of user information. Hacker can use backdoor channels to get control of victim's resources and utilize it as zombie to launch DDoS attack. This attack targets the confidentiality and availability of cloud users.

E . DOS Attacks

The attacker exploits zombies for sending a large number of network packets to overwhelm the available resources. Consequently, legitimate users are unable to access the services offered over the Internet. In cloud environment, the attacker may send huge number of requests through zombies to access VMs thus disabling their availability to legitimate users which is called DoS attack This attack targets the availability of cloud resources.

IV TYPES OF INTRUSION DETECTION SYSTEM

A. Network based Intrusion Detection System

Network intrusion detection system monitors and analyzes network traffic by reading individual packets through network layer and transport layer. It searches for any suspicious activity or network based attack such as Denial of Service (DoS) attack, port scans etc. Once an abnormal behavior in network traffic is identified, alert can be sent to system administrator. Most of the commercial IDSs are based on the NIDS such as Snort, Tcpdump and Natural flight Recorder (Mehmood, Habiba, et al., 2013). These are well known for general sized networks and convenient for implementation to detect intrusions. However, (Kumar, & Hanumanthappa, 2012) discussed the main issues of Snort IDS when integrating with distributed computing environment. To overcome the issues, they introduced new approach for handling these issues. For virtual network systems, multi phase distributed vulnerability detection and measurement technique has been proposed to detect DDoS attack (Chung, Khatkar, et al. 2013). It has detected attacks based on attack graph by analyzing network traffic flowing through virtual machines. It has significantly improved attack detection and mitigates attack consequences.

B. Host based Intrusion Detection System

Host based intrusion detection system monitors the individual host or device on the network by analyzing any change in the activity performed by host and events occurring within that host. It looks at every activity of host by checking application logs, system calls, and file-system modifications, inbound and outbound packets to and from host. If any suspicious activity is found, an alert is generated and sent to administrator to protect the system from malicious attack. Since majority of sectors prefer HIDS also after NIDS which are mainly based on the log file analysis of system. A model of HIDS has been developed based on log file analysis of Microsoft Windows XP operating system. It detects intrusions by matching predefined pattern with the logs of operating system (Ali, & Len, 2011)

C. Distributed based Intrusion Detection System

Distributed IDS (DIDS) also known as hybrid IDS, consists of two or more detection methods or systems i.e., NIDS, HIDS etc. This type of system is deployed over large distributed network like cloud computing so as all entities can communicate with each other and with network monitor such as central server In this way, all hosts deployed over network collect system information and send it to central server by converting it into standard format

D. VMM/Hypervisor based Intrusion Detection System

Hypervisor provides a platform for communication among VMs. Hypervisor based IDSs is deployed at the hypervisor layer. It helps in analysis of available information for detection of anomalous activities. The information is based on communication at various levels like communication between VM and hypervisor, between VMs and communication within the hypervisor based virtual network.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 09, September-2020

V METHODOLOGY

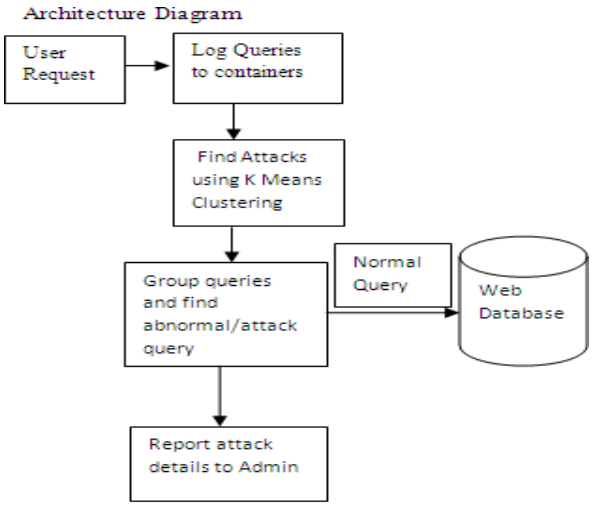


Fig 1 Architecture Diagram

Building normality model

A static testing website is deployed using the Content Management System assign each user session into a container, each container is assigned per each new IP address of the client. The container will log all the Web requests and SQL Queries executed by client. Deterministic Mapping and the Empty Query Set Mapping patterns are discovered from training sessions.

Attack scenarios

This system is effective and can reduce the computation time at detecting the following types of attacks:

A. Injection Attack

This SQL injection attack changes the structure of the SQL queries. It would generate SQL queries in a different structure that could be detected as a deviation from the SQL query structure that would normally follow such a web request.

Ex-

Original query=select * from admin where uid="1";

Suspicious query=select * from admin where uid="„ OR 1=1;--„

Here, original query is passed and suspicious query is blocked. Word-list contains the tokens of sql-query strings.

„O"-Original query

„S"-Suspecious query

Ex- („O") select * from admin where uid = „1„;

("S") select * from admin where uid = „ „ OR 1=1;--„

("O") select * from admin where uid ="1" && pwd ="abc";

(„S") select * from admin where uid = „ „ OR 1=1;--„

B. URL Manipulation

A Client manually adjusts the parameters of its request by maintaining the URL's syntax but altering its semantic meaning. This system can easily capture the attack because it would generate SQL queries that actually not match with the training sessions. *Privilege Escalation Attack*

A privilege escalation attack is a type of network intrusion that takes advantage of programming errors or design flaws to grant the attacker elevated access to the network and its associated data and applications.

Let's consider that the dynamic website serves both regular users and administrators. For a regular user, the web request Ru will trigger the set of SQL queries Qu, and for an administrator, the request Ra will trigger the set of admin level queries Qa. Now an attacker logs into the webserver as a normal user, upgrades his privileges and act as administrator to trigger admin queries to obtain administrators data.

This approach can detect this type of attack since the DB query Qa does not match with the request Ru according to the proposed mapping model.

C. Session Hijacking

Session Hijacking is an attack by which the hacker steals this user's session identifier and then sends this session identifier as their own to the server and tricks the server into thinking they are that user. By hijacking other user sessions, the attacker can eavesdrop, send spoofed replies, and/or drop user requests. A session- hijacking attack can be further categorized as a Spoofing/Man-in-the-Middle attack, an Exfiltration Attack, a Denial-of-Service/Packet Drop attack, or a Replay attack.

According to the mapping model, if found the same session identifier from two different IP would be treated as hijacked session

D. Mapping models

1. Finding deterministic mapping queries

Deterministic Mapping is the most common and perfectly matched pattern. Web request rm appears in all traffic with the SQL queries set Qn. If Qn is present in the session traffic without the corresponding rm is classified as intrusion [1].

2. Finding Empty Query Set

In special cases, the SQL query set may be the empty set. This implies that the web request neither causes

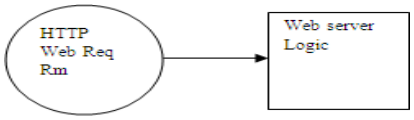


Fig 4 Empty Query Set

nor generates any database queries. For Example, when a web request for retrieving an image GIF file from the same webserver is made, a mapping relationship does not exist because only the web requests are observed. This type of mapping is called rm->O;. During the testing phase, keep these web requests together in the set EQS

3. No Matched Request

In Some cases, DB queries cannot match up with any web requests, therefore these queries are kept in No Match Request. During testing phase any query within the set No Match Request is considered as legitimate. If found anything abnormal then it is treated as attack.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 09, September-2020

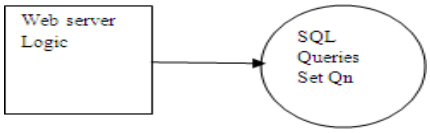


Fig 5 No Matched Request

4. Non Deterministic Mapping

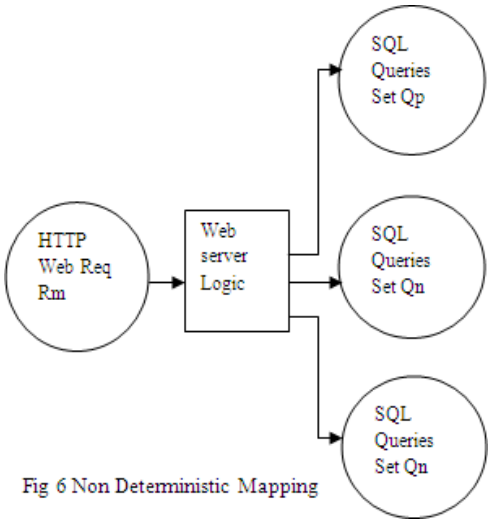


Fig 6 Non Deterministic Mapping

This happens only within dynamic websites, such as blogs or forum sites. Because the same web request may result in different SQL query sets based on input parameters Each time that the same type of web request arrives, it always matches up with one (and only one) of the query sets The mapping model is Rm € Qi (Qi € {Qp, Qr, Qq}).

Clustering Algorithm for Attack Classification

The algorithm accepts two inputs. The data(container) itself, and "k", the number of clusters. The output is k clusters with input data partitioned among them.
The aim of K-means (or clustering) is this: To group the items into k clusters such that all items in same cluster are as similar to each other as possible. And items not in same cluster are as different as possible. The distance measures are used to calculate similarity and dissimilarity. One of the important concept in K-means is that of centroid. Each cluster has a centroid. Consider it as the point that is most representative of the cluster. Equivalently, centroid is point that is the "center" of a cluster.
The output of the clustering is given as input to above modules to get optimized result for detecting intrusion in website.

K means algorithm for attack classification

1. Initialize logged queries for container
2. Randomly choose k queries and make them as initial centroids.
3. For each point, find the nearest centroid and assign the point to the cluster associated with the nearest centroid.
4. Update the centroid of each cluster based on the items in that cluster. Typically, the new centroid will be the average of all points in the cluster.
5. Repeats steps 2 and 3, till no point queries are clustered.

VI PERFORMANCE EVALUATION

To evaluate the output for this system, different attacks have been analyzed, as discussed in Section 4.2, by implementing real spatial datasets by developing online portal website for real-estate agencies and hosted in the cloud. The datasets given are 50, 150, 250 and the time required for existing and proposed method is also shown below in the table as such that the performance of the proposed system shows the low value of time taken. This system then compare the cost of the algorithms with respect for queries with influence scores. The below table shows the cost of the algorithms by varying the number k of requested results by retrieving time

TABLE I

No of Data's	Normal Time	Clustered time
50	.97 sec	.85 sec
150	1.25 sec	1.12 sec
250	1.45	1.23 sec

Fig. 7 Comparison between Normal and Cluster classification based on time

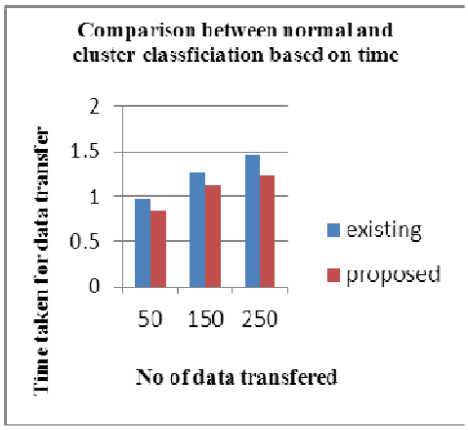
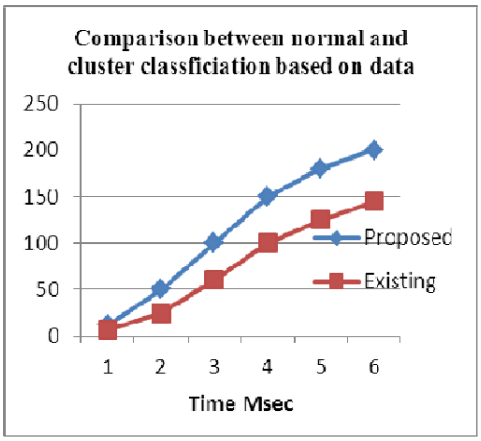


Fig. 8 Comparison between Normal and Cluster classification based on data



International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 09, September-2020

VII CONCLUSION

This paper presented an Intrusion detection system that builds models of normal behavior for Cloud based applications from both front-end web (HTTP) requests and back-end database (SQL) queries. Unlike previous approaches that correlated or summarized alerts generated by independent IDSs, This approach forms container- based IDS with multiple input streams to produce alerts. K Means Clustering Algorithm is used efficiently to classify attacks from logged Queries. We have shown that such correlation of input streams provides a better characterization of the system for anomaly detection because the intrusion sensor has a more precise normality model that detects a wider range of threats and reduce classification time largely.

VIII REFERENCES

[1] Distributed Intrusion Detection System for Cloud Environments based on Data Mining techniques , Procedia Computer Science 127 (2018) 35–41
[2] Intrusion Detection System for Cloud Computing, International Journal of Scientific & Technology Research Volume 1, Issue 4, May 2012
[3] Meixing Le, Angelos Stavrou, Brent ByungHoon Kang," DoubleGuard: Detecting Intrusions in Multitier Web Applications", IEEE Transactions On Dependable and Secure Computing , VOL. 9, NO. 4, JULY/AUGUST 2012
[4] Y. Shin, L. Williams, and T. Xie, "SQLUnitgen: Test Case Generation for SQL Injection Detection," technical report, Dept. of Computer Science, North Carolina State Univ., 2006.
[5] C. Anley, "Advanced Sql Injection in Sql Server Applications," technical report, Next Generation Security Software, Ltd., 2002.
[6] "Common Vulnerabilities and Exposures," http://www.cve.mitre. org/, 2011.
[7] "Five Common Web Application Vulnerabilities,"http://www.symantec.com/connect/articles/five - common-web-applicationvulnerabilities, 2011.
[8] "A Review of Anomaly based intrusion detection system"International Journal of Computer Applications(0975-8887) volume 28-No.7,August 2011
[9] Intrusion Recovery for Database-backed Web Applications, Ramesh Chandra, Taesoo Kim. ACM 2011
[10] https://www.uptycs.com/blog/intrusion-detection-in-cloud-computing