Analysis of Intrusion Detection System for Cloud Environment

[1] Saurabh P.Taley , [2] J.J.Shah

[1] CSE(ESC-IV Sem) , G.H.Raisoni College of Engineering, Nagpur, Maharastra, India

[2] Assistant prof. IT DEPARTMENT, G.H.Raisoni College of Engineering, Nagpur, Maharastra, India

Abstract - Cloud computing provides large scale computing resource to each customers. Cloud systems can be threatened by numerous attacks as cloud provides services to no trustworthy system. Cloud needs to contain intrusion detection system for protecting system against threads. If IDS is having stronger security using more rules or patterns then it need much more computing resources. Current cloud monitoring systems can rely on signature-based and supervised-learning-based detection methods to check out attacks and anomalies. Propose work introduce UCAD, an Unsupervised Cloud Anomaly Detection for knowledge-independent detection of anomalous traffic. UCAD uses a novel clustering technique based on Partition based clustering to identify clusters and outliers in multiple low-dimensional spaces. The evidence of traffic structure provided by these multiple clustering is then combined to produce an abnormality ranking of traffic flows, using a correlation-distance-based approach.

***Keywords* - Intrusion Detection System, Anomaly detection, Clustering.**

1. Introduction

Cloud computing has evolved through a number of Implementations. Moving data into the cloud provides great convenience to users. Cloud computing is a collection of all resources to enable resource sharing in terms of scalable infrastructures, middleware and application development platforms, and value-added business applications. The characteristics of cloud computing includes: virtual, scalable, efficient, and flexible. In cloud computing, three kinds of services are provided: Software as a Service (SaaS) systems, Infrastructure as a Service (IaaS) providers, and Platform as a Service (PaaS). In SaaS, systems offer complete online applications that can be directly executed by their users. In IaaS, providers allow their customers to have access to entire virtual machines; and in SaaS, it offers development and deployment tools, languages and APIs used to build, deploy and run applications in the cloud. The virtual environment lets users use computing power ,which far exceeds that contained in their physical worlds. These services in cloud computing may easily expose to the risk of security attacks. Within the cloud computing, security issues, such as confidentiality, integrity and availability (CIA) are the most important security considerations. Denial-of-service (DoS) attack and distributed denial-of-attack (DDoS) are other kinds of attacks that cause the targeted system or network unusable [2]. Therefore, if the cloud computing framework suffers from these kinds of attacks, the service providers and users could not use the services. Intrusion detection system (IDS) is a practical solution to resist these kinds of attacks. However, if IDS is deployed in each cloud computing region, but without any cooperation and communication, IDS may easily suffers from single point of failure attack. Obviously, the abilities of intrusion detection and response are decreased significantly. Thus, the cloud environment could not support services continually. In order to protect the cloud environment from DoS or DDoS attacks, the proposed paper launches an idea of federation defense in the cloud computing. Based on this concept, IDS system is deployed in each cloud computing region. These IDSs will cooperate with each other by exchanging alerts to reduce the impact of the DoS attack. Within this framework, Snort based IDS is implemented and three modules are plug-in into the system. These modules are block, communication and cooperation modules. Clustering is a process of labeling data and assigning that data into groups of similar objects [2]. Each group is called as cluster. It consists of members from the same cluster that are similar and members from the different clusters that are different from each other.

2. Literature Survey

In [2.1] the proposed method is based on K-means clustering, which is a typical clustering algorithm. K-means is one of the simplest unsupervised learning clustering algorithms. Its procedure follows an easy way to classify a given data set through a certain number of k clusters that are fixed a priori. The system include the hybrid approach for intrusion detection. It consists of feature selection, filtering, clustering, divide and merge, clustering ensemble and normal and intrusion detection. Feature selection selects the important attributes from the data set. A filter method helps in reducing noise and outliers on the data set. Divide and merge helps in calculating the k number of the cluster centroids. By the more accurate method of finding initial k clustering centers, the intrusion detection model with clustering ensemble is presented to achieve high accuracy and detection rate as well as very low false alarm rate. Hence a hybrid data mining approach for intrusion detection system is proposed in this paper. The main research method is clustering analysis with the aim to achieve high detection rate and very low or no false alarm rate [2].

In [2.2] this work present a partition based algorithm for outlier detection over data stream, where we solve the problem of outlier detection using LOF (Local Outlier Factor Algorithm) algorithm over the sparse and dense regions separately. Clustering algorithm is incorporated for the process of efficient partitioning of the data stream chunks [11]. Through number of experiments it is shown that, by applying LOF on these partitions separately with variation in case of sparse regions and by dividing the stream in chunks, we can reduce memory consumption, number of nearest neighbour searches, number of rechability distance computation, and LOF computations for every element.

In [2.3] these algorithms detect outliers according to the distance between pairs of objects Distance-based outlier detection methods only take into account the distance between data objects, while density-based methods focus on the number of adjacent points. In anomaly intrusion detection, if one object of data set is far away from other objects and there are few objects around it, users consider it as outlier and it is believed to be a attack. So user can figure out gravity of every data object as the degree of its isolation. Considering intrusion detection data set is large-scale and it takes long time to compute every object's gravity between others, this paper makes cluster analysis for data set firstly, then figures out the degree of gravity between an object and a class by formally simulating a gravity function, and treat calculated value as outliers' discriminating criterion. Because the number of cluster is very small compared to the number of objects in intrusion detection data set, the time efficiency is better and time complexity of this algorithm is analyzed detail. The test results indicate this algorithm has perfect detection performance, especially for attacks. and the time complexity is nearly linear[12] with the size of dataset
In [2.4] System base its intrusion detection methods on the Perron Frobenius theorem.“ Perron Frobenius theorem asserts that a real square matrix with positive entries has a unique largest real eigenvalue and that the corresponding eigenvector has strictly positive components.” This algorithm using the Perron Frobenius theorem gives dynamic evaluation: with each successive iteration, user asks if the security score remains stable or will be different [4].

In [2.5] an anomaly detection algorithm using improved hierarchical clustering (ADIHC). There are two major advantages of ADIHC: firstly, It propose the novel density-based pruning algorithm. It not only optimize and compress the searching space, but also filter some noise. The relevant experiments shows the density-based pruning method improve the overall performance of detection after filtering some noise. Secondly, with the help of the improved hierarchy clustering structure, normality profiles can be updated at any time. The operation of inserting and deleting can be carried out as the same time as detection. Compared with traditional clustering algorithms, ADIHC has lower false alarm rate and higher detection rate. The superior performance of detection is mainly due to the high accuracy of normality profiles and the anti noise capability of ADIHC itself [15].

In [2.6] the proposed to perform joint learning for detecting intrusion when intruders do not carry any wireless devices (e.g., intrusion to corporate assets or people trapped in a fire building). joint intrusion learning approach combines the detection power of complementary intrusion indicators and has the capability to detect different intrusion events in wireless environments. In particular system utilized the Received Signal Strength (RSS) from the existing wireless infrastructure and exploited to use the changes of RSS causedby intrusions for diagnosing the presence of intrusions. profiled environmental uncertainties through data cleansing and intrusion pattern derivation. the grid-based clustering over K-neighborhood (GREEK) algorithm, which captures the declustering effect in intrusion indicators when intrusions are present[12].

In [2.7] Principal component analysis (PCA) neural network module which preprocess the input stream to reduce the dimension of the input space; rule creation and management (RCM) module which create and extract rules according to the achieved resulting clusters by the ECMm algorithm and the two level Fuzzy Inference System classifier(FIS) module. The RCM module connect FIS module with ECMm, in order to partition the input space. The results from ECMm are transitioned into the RCM module. The rules are extracted using information of each cluster, thus each cluster represents a kind of rule[16] . The ECMm algorithm will produce new cluster with which new classifiers are trained and added into the FIS module.

3. Comparison of Techniques

Comparison of techniques depends on various parameters, such as Algorithms used by techniques, its Advantages depends on their performance, various advantages of techniques depend on parameter like speed, filters ,sound ,time to find solutions. All this comparison is described by table below which shows the various algorithms, techniques and advantages of it.

3.1 Comparison of Clustering Algorithm

Table 1 – Comparison of clustering algorithm

NAME OF TECHNIQUES	ADVANTAGES	ALGORITHM USED
Model based clustering [15]	I. Find more suitable solution even more critical	Expectation maximization algorithm (EM algorithm)
Partition based clustering [12]	I.LOF(local outlier factor) II. LRD (local reachable distance	local outlier factor algorithm with K-nearest neighbor
Partitioning based clustering[2]	I. High detection rate II .high Accuracy	K-means algorithm
Density based clustering[14]	I.Hundle noise II.one scan	Gravity based anomaly intrusion detection algorithm
Partitioning based clustering[2]	I. Divide and merge II. A filter method	K-means algorithm
Grid based clustering [11]	I.Provide number of alerts	Grid-based clustering over K-neighborhood algorithm(GREEK)
Density based clustering[5]	I. Detected with accurate signatures(high security)	Evidence Accumulation for Ranking Outliers

3.2 Proposed Architecture

Proposed approach relies on robust clustering algorithms to detect both well-known as well as completely unknown attacks, and to automatically produce easy-to-interpret signatures to characterize them, both in an on-line basis. The analysis is performed on packet-level traffic, captured in consecutive time slots of fixed length. IP flows are additionally aggregated at different flow levels . These include: source IPs, destination IPs, source Network Prefixes, destination Network Prefixes, and traffic per Time Slot. The complete detection and characterization algorithm runs in three successive stages. The first step consists in detecting an anomalous time slot where an attack might be hidden. The unsupervised detection and characterization algorithm begins in the second stage, using as input the set of IP flows captured in the flagged time slot. The method uses robust clustering techniques based on K-MEANS to blindly extract the suspicious flows that compose the attack. In the third stage, the evidence of traffic structure provided by the clustering algorithms is used to produce filtering rules that characterize the detected attack and simplify its analysis. The characterization of an attack can be a hard and Time-consuming task, particularly when dealing with unknown attacks. Even expert operators can be quickly overwhelmed if simple and easy-to-interpret information is not provided to prioritize the time spent in the analysis. To alleviate this issue, the most relevant filtering rules are combined into a new traffic signature that characterizes the attack in simple terms. Algorithm automatically produces new signature without any previous data about traffic or knowledge about the attack.

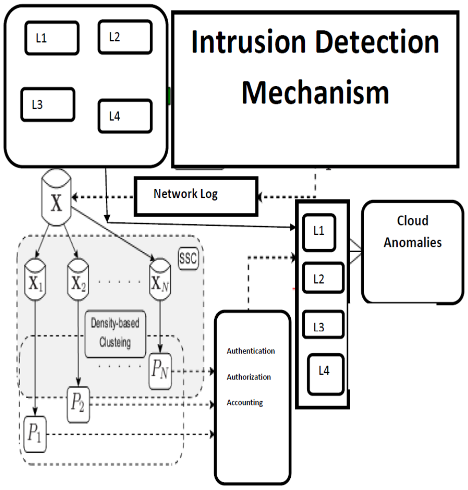

Fig. 1 Architecture of Proposed System

4. Conclusions

IDS methods leads to effective resources usages by applying differentiated level of security strength to users based on the degree of anomaly. Through the cloud computing it is possible to judge all users and administrators as potential attacker and apply strong security policy to all traffic, but it is not efficient at all. If any security hazards occur, economic damages are unavoidable. proposed work Unsupervised cloud Anomaly Detection UCAD verified the effectiveness of the system detect real single source-destination and distributed attacks in real traffic all in a completely blind fashion, without assuming any particular traffic model, clustering parameters, or even clusters structure beyond a basic definition of what an anomaly. So it's effective way of detecting attacks in cloud environment.

References

[1] Zhen Chen*, Fuye Han, Junwei Cao, Xin Jiang, and Shuo Chen "Cloud Computing-Based Forensic Analysis for Collaborative Network Security Management System" TSINGHUA SCIENCE AND TECHNOLOGY ,2013.

[2] Kapil Wankhade, Sadia Patka, Ravindra Thool " An Efficient Approach for Intrusion Detection Using Data Mining Methods" 2013 IEEE

[3] En Niari Saad, Khalil El Mahdi, Mostapha Zbakh "Cloud Computing Architectures Based IDS" 2012 IEEE.

[4] Amira Bradai and Hossam Afifi "Enforcing Trust-based Intrusion Detection in Cloud Computing Using Algebraic Methods" 2012.

[5] mld Khoudali, Karim Benzidane and Abderrahim Sekkaki "Inter-VM packet inspection in Cloud Computing " The 5th International Conference on Communications, Computers and Appllication 2012.

[6] Massimo Ficc o, Massimiliano Rak, and Beniamino Di Martino "An Intrusion Detection Framework for Supporting SLA Assessment in Cloud Computing" 2 0 12 I EEE.

[7] Matthias Gander, Basel Katt, Michael Felderer, Adrian Tolbaru, Ruth Breu, Alessandro Moschitti "Anomaly Detection in the Cloud: Detecting Security Incidents via Machine Learning" 2012.

[8] Hisham A.Kholidy,Fabrizio Baiardi "DCDIDP:A Distributed, Collaborative, and Data-driven Intrusion Detection and Prevention Framework for Cloud Computing Environments" 2012 IEEE.

[9] Pedro Casas, Johan Mazel and Philippe Owezarski "UNADA: Unsupervised Network Anomaly Detection using Sub-Space Outliers Ranking" 2011.

[10] Jun-Ho Lee,Min-Woo Park,Jung-Ho Eom, and Tai-Myoung Chung"Multi-level Intrusion Detection System and Log Management in Cloud Computing" 2011.

[11] Jie Yang, Yong Ge, Hui Xiong, Yingying Chen, Hongbo Liu "Performing Joint Learning for Passive Intrusion Detection in Pervasive Wireless Environments " 2010.

[12] Manzoor Elahi,Kun Li,Wasif Nisar,Xinjie Lv,Hongan Wang "Detection of Local Outlier Over Dynamic Data Streams using Efficient Partitioning Method" 2009

[13] Baoyi Wang, Ranran Jin, Shaomin Zhang, Xiaomin Zhao "Research on Gravity-based Anomaly Intrusion Detection Algorithm" 2009.

[14] HU Liang, REN Wei-wu, REN Fei "Anomaly Detection using Improved Hierarchy Clustering" 2009.

[15] Yu-Ping Zhou, Jian-An Fang , Yu-Ping Zhou "Research on Neuro-Fuzzy Inference System in Hierarchical Intrusion Detection"2009.

.