International Journal of Information Security and Privacy
Volume 14 • Issue 1 • January-March 2020

An Improved Intrusion Detection System to Preserve Security in Cloud Environment

Partha Ghosh, Netaji Subhash Engineering College, MAKAUT, Kolkata, India
Sumit Biswas, Tata Consultancy Services Ltd, Mumbai, India
Shivam Shakti, Netaji Subhash Engineering College, MAKAUT, Kolkata, India
Santanu Phadikar, Maulana Abul Kalam Azad University of Technology, Kolkata, India

ABSTRACT

Cloud computing, also known as on-demand computing, provides different kinds of services for the users. As the name suggests, its increasing demand makes it prone to various intruders affecting the privacy and integrity of the data stored in the cloud. To cope with this situation, intrusion detection systems (IDS) are implemented in the cloud. An effective IDS constitutes of less time-consuming algorithm with less space complexity and higher accuracy. To do so, the number of features are reduced while maintaining minimal loss of information. In this paper, the authors have proposed a model by which the features are selected on the basis of mutual information gain among correlated features. To achieve this, they first group the features according to the correlativity. Then from each group, the features with the highest mutual information gain in their respective groups are selected. This led them to a reduced feature set which provides quick learning and thus produces a better IDS that would secure the data in the cloud.

KEYWORDS

Cloud Computing, Core Cluster, Feature Selection, Intrusion Detection System (IDS), Mutual Information (MI)

INTRODUCTION

Cloud computing is a widespread term for the transportation of hosted services using the Internet. Cloud computing has evolved as one of the most vital dimension of the modern software industry by making a transition from computing-as-a-product to computing-as-a-service (Murugesan, 2011). Instead of setting up a physical infrastructure, Cloud allows us to have the luxury of using applications, software, platforms etc. as a service and one has to pay only for the resources he consumes (Singh & Jangwal, 2012). Since in a Cloud Environment data arrives from different heterogeneous sources therefore understanding the associative vulnerabilities is the foremost job to do (Grobauer, Walloschek, & Stöcker, 2011) and then, to provide a way to maintain the integrity, confidentiality and availability of the incoming and outgoing data. Hamlen et al. (Hamlen, Kantarcioglu, Khan, & Thuraisingham, 2010) in their work have discussed the various security issues of the Cloud. IDS is one such solution that provides data security to the Cloud Environment. Based on deployment, IDS have two models, Host Based IDS(HIDS) and Network Based IDS(NIDS). HIDS attempts to recognize unauthorized, abnormal behaviors on a specific device (Hu, 2010). HIDS uses both Anomaly Based and Misuse Based Detection Techniques and plays a very compliant role in identifying, logging records and alerting the admin if there is any security issue. Whereas NIDS completely works on Network

DOI: 10.4018/IJISP.2020010105

traffic. It captures Ethernet Packets and scans it in real time to decide whether it is an attack or not (Mukherjee, Heberlein, & Levitt, 1994). The number of unnecessary generated alerts in Anomaly Based IDS which causes high false alarm can be reduced as demonstrated by Hacini et al. (Salima Hacini, Zahia Guessoum, 2013) .

As the network traffic is huge in size so the analysis of packets in real time is too time-consuming phenomenon, hence for better performance of IDS it is incorporated with various data mining algorithms extensively (Yanjie, 2015). For further enhancement in the performance pre-processing of data becomes inevitable which reduces dimensions quite significantly (Said, Stirling, Federolf, & Barker, 2011). Feature Selection is one of the most widely used pre-processing technique which eliminates irrelevant and homogeneous features from a given feature set (Mladeni, 2006). Another pre-processing technique is Clustering which helps to eliminate outliers, noise and group similar kind of objects. Objects can be either instances or features (Kryszkiewicz & Skonieczny, 2005). For the experimental purpose the authors have used NSL-KDD dataset for training and testing purpose. In this paper, initially authors have designed a fully connected weighted graph of features, where each node represents a feature. Then Core Clusters are created by removing inconsistent edges. Later, the relevant features which have high Mutual Information Values, are selected from each core in order to get the Relevant Feature Set (RFS). Using the above mentioned methods the authors have proposed an Anomaly Based Intrusion Detection System.

RELATED WORK

Cloud has been an inseparable part of modern day technology because of its fascinating storage and computing capability. Therefore various conventional services such as Messaging Services, Social Networking Services are shifting towards Cloud Platform. Shawish and Salama (Shawish & Salama, 2014) gave a overview of Cloud's anatomy, characteristics and architecture. They also covered a detailed comparison between Cloud Based Services and Existing Services. Though Cloud is very flexible but it is quite vulnerable to various kinds of attacks. To overcome all these data security issues various IDSs are required. A brief introduction to IDS was proposed by Mohamed et al. (A. Mohamed, Idris, & Shanmugum, 2012). In their work they reviewed IDS, pointed out those issues that appeared during implementation of IDS and the restrictions in the research field in IDS. Kumar et al. (B. S. Kumar et al., 2001) provided an in depth description of IDS model and the various types of intrusion in the system and their corresponding prevention techniques. Lombardi and Di Pietro (Lombardi & Di Pietro, 2011) proved how Virtualization can be implemented to increase security in Cloud. They proposed a novel architecture, Advanced Cloud Protection System(ACPS) that can fruitfully audit the integrity of the Cloud Environment. Denning (Denning, 1987) developed a general purpose IDS framework which was system as well as environment independent and consolidated the fact that, security breaches can be identified by monitoring system's log of unusual patterns. Kholidy and Baiardi (Kholidy & Baiardi, 2012) outlined a framework to work out the inadequacy of IDS model. They incorporated both Knowledge-Based and Behavior-Based techniques to improve the overall attack handling capability. Later on, for the betterment of IDS performance in Cloud, several data mining techniques were also introduced. Lee and Stolfo (Lee & Stolfo, 2000) suggested a model which was based on data mining algorithms. The model used Classification, Meta-Classification, Association and Frequent Rule to generate frequent pattern from audit log to detect anomalies. The result displays that the model is as good as those systems which were manual knowledge approach driven. Mohamed et al. (S. Mohamed, Mohamed, & Mokhtar, 2017) proposed an IDS model using a hybrid approach of K-Means and Sequential Minimal Optimization (SMO) Classification. They apply the approach on NSL-KDD dataset and the result shows that it brings down the false alarm rate quite magnificently and achieves higher accuracy. Few Denials of Service(DoS) attacks can bypass both the application and operating system layer which imposes serious threats. That's why Tao et al. (Tao, Yang, Peng, & Li, n.d.)proposed a HIDS which shows better detection rate of DoS