Cloud Computing Security aspects: Threats, Countermeasures and Intrusion Detection using Support Vector Machine

Sarvottam Dixit [1], Gousiya Hussain[2]

[1]Mewar University, Chittorgarh, India
[2]Mewar University, Chittorgarh, India.

Article History Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract. The Cloud Computing concept has witnessed immense growth, and the rationale behind this increased interest in the emerging paradigm is the cost-effective transmission, storage and intensive computation. The idea is to render remote storage and data analysis capability to the end-user using shared computing resources, thereby bringing down the total cost incurred for an individual. However, consumers are still reluctant to adopt this technology due to associated security and privacy concerns. This manuscript presents a comprehensive description of various threats and technical security concerns with respect to cloud computing. Additionally, support vector machine was used to detect intrusion on the NSL-KDD dataset designed by DARPA. The accuracy of the algorithm for DoS and probe attack were analyzed, and the results are represented in the form of confusion matrices. Among the 41 attributes of the NSL-KDD dataset, only 24 significant attributes are selected. We observed that our model could predict a DoS attack with an accuracy of 96% and a probe attack with an accuracy of 84%.

Keywords: Security, Threats, cloud computing, Intrusion Detection, SVM, NSL-KDD

1. Introduction

The emergence of cloud computing as a mainstream solution to big data processing has revolutionized the digital world and lead to *remote* and *enmasse* computing service provision by third-party service providers (Hayes, 2008) (Prajapati & Shah, 2020).Due to low-cost and flexible computing services (like infrastructure, platform and software) and computing resources (like hardware, CPU and storage etc.) provided by cloud service providers, the consumers can get access to services that otherwise would be inaccessible and too expensive(Campanile et al., 2020). Multiple definitions of cloud computing have surfaced over the years, leading to confusion among people about what exactly cloud computing is?

The US NIST, in its special publication (SP) 800-145 (Mell & Grance, 2011),has comprehensively framed the Cloud Computing definition in a generic way such that it includes most of the important te38rms related to this technology. It defines Cloud Computing as *"a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"*.

The rationale behind increasing interest in the emerging paradigm is a cost-effective transmission, storage and intensive computation(Mthunzi et al., 2020). The idea is to render remote storage and data analysis capability to the end-user using shared computing resources, thereby bringing down the total cost incurred for an individual(Sun, 2020). The economic interest in cloud computing for a cloud user is that he has to pay only for what he actually uses at any instance of time and from anywhere (Shuaib et al., 2019). With the advent of recent cloud-based technologies like IoT, Big Data, 5G, SDN and NFV, exponential growth is expected in the popularity and success of cloud-based opportunities(Lui et al., 2011)(F. Masoodi et al., 2019). Despite numerous benefits and advantages, consumers are reluctant to adopt this paradigm due to associated security and privacy concerns(Kumar & Goyal, 2019). The inherent vulnerabilities, potential risks, trust among consumers and service providers and openness in sharing technologies makes cloud computing susceptible to various threats and attacks. Intrusion is one of the most common threats, and in recent times has emerged as an important field for Cloud Security (Siddiqui, Alam, Khan, et al., 2019). An intrusion detection system (IDS) is a system designed to detect undesirable attacks to change, modify or disable system resources, mainly through the internet. IDS analyses the network traffic to check for activities that can jeopardize the services like confidentiality, integrity, or availability. IDS generates an alarm to indicate the place where an attack is present. The intrusion detection system's primary challenge includes the problem of false alarm and inadequate support for real-time response to the attacks. It is an important tool for network administrators because, without such a device, it would not be possible to analyse a hefty amount of data travelling over the network.

2. ***Cloud Security Requirements****:* The security services *confidentiality, integrity, and availability discussed in Table 1 are considered fundamental requirements for cloud computing. NIST* has defined these security services as a requirement for the cloud.(Lui et al., 2011)(Hashizume et al., 2011).

Table 1.Security requirements for cloud computing

Requirement	**Requirement Description**
Confidentiality	Confidentiality refers to securing users' data against any kind of disclosure to unauthorized users, systems, or processes. Weak identification or vulnerabilities in applications can be exploited to gain unauthorized access (Mohammad Ubaidullah Bokhari et al., 2014).
Availability	Availability refers to facilitating the access and use of data and system to authorized entities (user, process or device) at the time of need. Availability refers to data and software and guarantees the availability of hardware as well, whenever required.
Integrity	Integrity refers to safeguarding data from any unauthorized change (modification, deletion, fabrication). Any change carried out to the data must be done by authorized users only (M U Bokhari & Alam, 2013).

3. ***Threats and countermeasures:*** Any event or a circumstance that has the potential of hampering or affecting an organization's operations, assets or individuals adversely is referred to as a threat. A host of potential threats discovered by Cloud Security Alliance (CSA) are discussed in Table 2 below. (Hamlen et al., 2010)(Jensen et al., 2009)(NIST, 2013)(F. S. Masoodi & Bokhari, 2019)

Table 2.Top threats and suggested countermeasures.

Top Threat	**Threats**	**Description**	**Countermeasures (suggested)**
1	Data breaches	Any unauthorized use of data that was accessed (stolen, leaked or released) with malicious intent.	Encryption
2	Weak identity, credential and Access management	Unauthorized users and access sensitive data can exploit weak identity, sensitive information like passwords and access management by pretending to be authorized users.	Multifactor Authentication
3	Insecure Application Programming Interfaces	Cloud APIs facilitate the development and management of cloud resources but, at the same time, expose the environment to the outside world that makes these APIs an entry point to the attackers.	Authentication and Access control and avoid reuse of API keys
4	System and Application Vulnerabilities	The implicit bugs or weakness present in the system and application may be exploited to threaten the security of the system	Patch management and constant vulnerability scanning
5	Account Hijacking	The attacker may steal the legitimate credentials and gain access to the cloud and compromise the cloud security services, which he is not supposed to access otherwise.	Two-tier authentication and least possible sharing of account credentials

6	Malicious Insiders	An insider employee (ex or current) with access to the system holding a grudge against the organization or partners may misuse the information and hamper the system in terms of security and privacy	Minimize access given to users and determine security breaches
7	Advanced Persistent Threats (APTs)	These kinds of threats arise due to sophisticated attacks that remain in the system for an extended period of times and adapt themselves to the security measures taken by the system	Proactive security measures and awareness.
8	Data Loss	Data loss may occur due to multiple reasons, including physical damage to the storage, accidental deletion, and malicious attacks. Periodic backups can help in restoring data.	Encryption
9	Insufficient Due Diligence	During the process of adopting cloud technology, the migration is done without a proper understanding of the new technology and security issues. Inadequate knowledge of the system may lead to potential security risks	Extensive due diligence before migration to the cloud
10	Abuse and Nefarious Use of Cloud Computing	The primary reason behind this kind of threat is a weak registration system that leads to unaccounted, mismanaged and weakly secured accounts that can launch attacks on a target system	Stern measures need to be taken during the registration and validation process.
11	Denial of service	An attempt to inhibit legitimate users from accessing cloud resources by attacking the target service with too many requests and slowdown the response time, resulting in little or zero access to genuine users.	Higher bandwidth, throttling and rate limit technologies.
12	Shared Technology Vulnerabilities	The primary idea behind the cloud computing concept is the sharing of resources among its consumers, making it vulnerable to attacks performed on shared resources.	Enforce strict security measures during installations and configuration.

4. Intrusion Detection in Cloud Computing Environment

One of the potential reason why cloud computing is vulnerable to cyber-attacks is its distributed and open structure(Patel et al., 2013). Intrusion detection and prevention systems play a vital role in early detection of such activities and even possible prevention of such attacks (Siddiqui, Alam, Shuaib, et al., 2019). Successful implementation of the intrusion detection system in the cloud computing environment requires considering various characteristics of cloud computing, including reliability, availability, quality of service, elasticity, and adaptability(Patel et al., 2013). In this work, we worked on the classification of malicious activities by using the Support Vector Machine (SVM) classification technique, and the dataset used is the benchmark NSL-KDD.

4.1 Methodology

In this section, the support vector machine-based intrusion detection technique results have been discussed using an interpreted object-oriented programming language, python. The dataset used for intrusion detection has been described, and the training and testing set from this data have been examined. Finally, the results are presented, and the performance of the detection method has been discussed.

4.2 Data Preprocessing

The focus here is to remove or substitute symbolic or non-numeric attributes, as they do not enhance accuracy in intrusion detection(Dash & Liu, 1997). In current research work, only 24 attributes were considered out of 41 attributes, which led to an increase in accuracy, besides reducing the execution time. The knowledge obtained from the model can be generalized by the algorithm if it works with all the attributes present in the dataset. Hence, it was important to choose only relevant attributes to recognize only the similar behavioral patterns present in the data could be avoided(Bamhdi et al., 2021). This removal of attributes is known as the reduction of dimensionality. Thus, pre-processing of data is important so that the overhead in processing the data can be minimized and also, important input features can be identified for intrusion detection

4.3 The Dataset

KDD dataset(Dhanabal & Shantharajah, 2015)is employed for training intrusion detection system. It is composed of different attack classes, viz: DOS, Probe R2L and U2R, as shown in Figure 1. The training data consists of around 67343 normal records, 45927 DOS entries and 11656 probe records extracted randomly from the NSL-KDD data set. Similarly, the testing data file consists of 9571 normal records, 8044 DOS entries and 2152 probe records (Abrar et al., 2020). The dataset is first trained on a sufficient number of attack entries and then tested to detect DOS and Probe attacks.

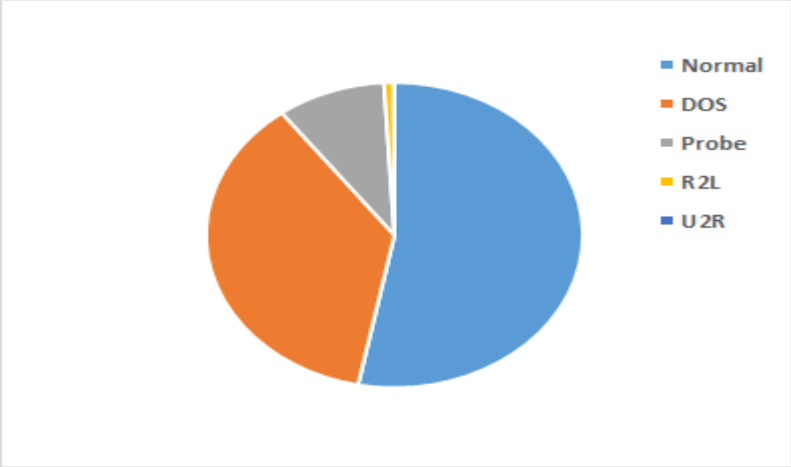

Figure 1: Normal & Attack class distribution

4.4 Results

Training of SVM was done on 125973 instances having 41 features. Similarly, testing was performed on 22,599 instances. The performance of the intrusion detection system is evaluated in terms of accuracy, precision, and recall.

Accuracy gives us the ratio of the total number of correct predictions, i.e. the sum of true positives (TP) and true negatives (TN), to the total number of values present in the dataset.

Accuracy = (TP +TN)/ (TP +T N+FP+FN)

Precision: The ratio of the correct number of positive predictions (TP) to the total number of positive predictions (TP + FP).

Precision= TP/ (TP + FP)

Recall/sensitivity: The ratio of the correct number of positive predictions (TP) to the total number of positive (P) values.

Recall = TP/P

The results have been presented in the form of a confusion metric (see Table 3) and Figure 4 gives the Graphical representation of results during the testing phase

Table 3: Confusion Matrices, Precision and Recall for DOS and PROBE attacks during Testing Phase

*a. Confusion matrix for **DOS Attack** Training (L) and Testing (R)*

	Positive	**Negative**
Positive	67236	107
Negative	338	67005

	Positive	**Negative**
Positive	6013	1445
Negative	313	9398

*b. Confusion matrix for **PROBE Attack** Training (L) and Testing (R)*

	Positive	**Negative**
Positive	67001	342
Negative	509	66834

	Positive	**Negative**
Positive	9328	383
Negative	428	1993

*c. Precision and Recall for **DOS attack***

Precision	**Recall**	**F1-Score**	**Support**
96%	81%	87%	7458

*d. Precision and Recall for **PROBE attack***

Precision	**Recall**	**F1-Score**	**Support**
84%	82%	83%	2421

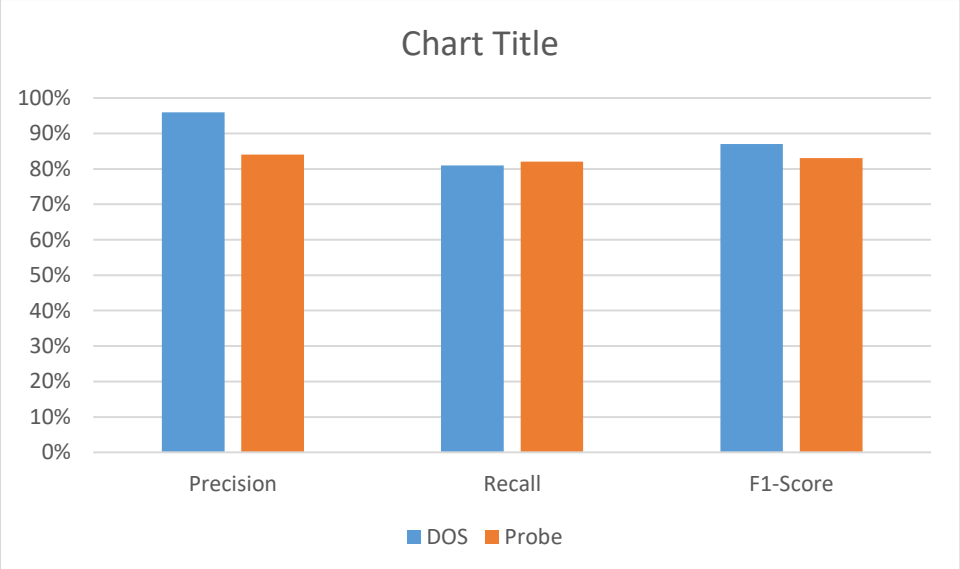

Figure 2: Graphical representation of results during the testing phase

5. Conclusion

Cloud-based systems have witnessed rapid development over the past decade. Even though the benefits easily outweigh the limitations, many practical issues need to be addressed. One such major concern is the associated security issues with cloud computing and is considered one of the fundamental reasons for inhibiting the implementation of cloud computing implementation. The study presented in this paper can help better understand

the various threats and security issues pertaining to cloud computing and used to design countermeasures to an existing set of threats. In this work, the intrusion was detected in the NSL-KDD dataset. The dataset was pre-processed by choosing 24 attributes based on their significance. Training and testing of data were carried out using SVM in which Dos and probe attacks were detected. The algorithm could detect Dos attack with 96% accuracy and probe attacks with 84% accuracy. Thus, SVM can be used effectively to detect intrusion in real-time networks.

References

1. Abrar, I., Ayub, Z., Masoodi, F., & Bamhdi, A. M. (2020). A Machine Learning Approach for Intrusion Detection System on NSL-KDD Dataset. *Proceedings - International Conference on Smart Electronics and Communication, ICOSEC 2020*, 919–924. https://doi.org/10.1109/ICOSEC49089.2020.9215232
2. Bamhdi, A. M., Abrar, I., & Masoodi, F. (2021). An ensemble based approach for effective intrusion detection using majority voting. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(2), 664–671. https://doi.org/10.12928/TELKOMNIKA.v19i2.18325
3. Bokhari, M U, & Alam, S. (2013). BSF-128: a new synchronous stream cipher design. *Proceeding of International Conference on Emerging Trends in Engineering and Technology*, 541–545.
4. Bokhari, Mohammad Ubaidullah, Alam, S., & Hasan, S. H. (2014). A Detailed Analysis of Grain family of Stream Ciphers. *International Journal of Computer Network & Information Security*, *6*(6).
5. Campanile, L., Iacono, M., Marrone, S., & Mastroianni, M. (2020). On Performance Evaluation of Security Monitoring in Multitenant Cloud Applications. *Electronic Notes in Theoretical Computer Science*, *353*, 107–127. https://doi.org/10.1016/j.entcs.2020.09.020
6. Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, *1*(3), 131–156. https://doi.org/10.3233/IDA-1997-1302
7. Dhanabal, L., & Shantharajah, S. P. (2015). A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, *4*(6), 446–452. https://doi.org/10.17148/IJARCCE.2015.4696
8. Hamlen, K., Kantarcioglu, M., Khan, L., & Thuraisingham, B. (2010). Security issues for cloud computing. *International Journal of Information Security and Privacy*, *4*(2), 36–48. https://doi.org/10.4018/jisp.2010040103
9. Hashizume, K., Rosado, D. G., Fernandez-Medina, E., & Fernandez, E. (2011). An analysis of security related issues in cloud computing. *Communications in Computer and Information Science*, *168 CCIS*, 180–190.
10. Hayes, B. (2008). Cloud Computing. *Communications of the ACM*, *51*(7), 9–11. https://doi.org/10.1145/1364782.1364786
11. Jensen, M., Schwenk, J., Gruschka, N., & Iacono, L. Lo. (2009). On technical security issues in cloud computing. *CLOUD 2009 - 2009 IEEE International Conference on Cloud Computing*, 109–116. https://doi.org/10.1109/CLOUD.2009.60
12. Kumar, R., & Goyal, R. (2019). On cloud security requirements, threats, vulnerabilities and countermeasures: A survey. *Computer Science Review*, *33*, 1–48. https://doi.org/10.1016/j.cosrev.2019.05.002
13. Lui, F., Tong, J., Mao, J., Bohn, R., Messina, J., Badger, L., & Leaf, D. (2011). NIST Cloud Computing Reference Architecture: Recommendations of NIST. *National Institute of Standard and Technology*, *NIST Speci*, 1–35.
14. Masoodi, F., Alam, S., & Siddiqui, S. T. (2019). Security and privacy threats, attacks and countermeasures in Internet of Things. *Int. J. Netw. Secur. Appl*, 67–77.
15. Masoodi, F. S., & Bokhari, M. U. (2019). Symmetric Algorithms I. *Emerging Security Algorithms and Techniques*, *January*, 79–95. https://doi.org/10.1201/9781351021708-6
16. Mell, P., & Grance, T. (2011). The NIST-National Institute of Standars and Technology- Definition of Cloud Computing. *NIST Special Publication 800-145*, 7.
17. Mthunzi, S. N., Benkhelifa, E., Bosakowski, T., Guegan, C. G., & Barhamgi, M. (2020). Cloud computing security taxonomy: From an atomistic to a holistic view. *Future Generation Computer Systems*, *107*, 620–644. https://doi.org/10.1016/j.future.2019.11.013
18. NIST. (2013). Glossary of key information security terms. In *NIST IR* (Vol. 7298, Issue Revision 2).
19. Patel, A., Taghavi, M., Bakhtiyari, K., & Celestino Júnior, J. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of Network and Computer Applications*, *36*(1), 25–41. https://doi.org/10.1016/j.jnca.2012.08.007
20. Prajapati, P., & Shah, P. (2020). A Review on Secure Data Deduplication: Cloud Storage Security Issue. *Journal of King Saud University - Computer and Information Sciences*, *xxxx*. https://doi.org/10.1016/j.jksuci.2020.10.021

21. Shuaib, M., Samad, A., Alam, S., & Siddiqui, S. T. (2019). Why Adopting Cloud Is Still a Challenge?—A Review on Issues and Challenges for Cloud Migration in Organizations. In *Advances in Intelligent Systems and Computing* (Vol. 904, pp. 387–399). https://doi.org/10.1007/978-981-13-5934-7_35
22. Siddiqui, S. T., Alam, S., Khan, Z. A., & Gupta, A. (2019). Cloud-based E-learning: using cloud computing platform for an effective e-learning. In *Smart Innovations in Communication and Computational Sciences* (pp. 335–346). Springer.
23. Siddiqui, S. T., Alam, S., Shuaib, M., & Gupta, A. (2019). Cloud Computing Security using Blockchain. *Journal of Emerging Technologies and Innovative Research*, *6*(6), 791–794.
24. Sun, P. J. (2020). Security and privacy protection in cloud computing: Discussions and challenges. *Journal of Network and Computer Applications*, *160*, 102642. https://doi.org/10.1016/j.jnca.2020.102642