International Research Journal of Commerce and Law
Volume 5 Issue 10, October 2018 ISSN: 2349-705X Impact Factor: 4.616
Journal Homepage: http://ijmr.net.in, Email: irjmss@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

Intrusion detection system for cloud based infrastructure using machine learning.

Prof. Shende Sachin Santosh,Prof. Madane Tai Abaso, Prof. Pansare Rajashree Balasaheb, Prof. Dhaygude Tejashri Mohan, Prof. Pandi M.

DattakalaShikshanSanstha "Dattakala Group of Institution" Swami-Chincholi, Daund, Pune, Maharashtra 413130. India.

Abstract: *This research addresses the burgeoning challenges in Cloud computing security through the proposition of an Intrusion Detection System (IDS) driven by Machine Learning (ML). Conducting a comprehensive literature review, the study elucidates the multifaceted landscape of security concerns in Cloud computing, spanning trust issues to sophisticated attack vectors. The proposed IDS integrates diverse ML algorithms, including Support Vector Machines and Neural Networks, strategically applied for effective intrusion detection in the Cloud. A mathematical foundation is established for the Support Vector Machine algorithm, showcasing its potential in threat detection through binary classification. The research delineates hardware and software requirements, emphasizing scalability and compatibility with prevalent ML frameworks. Despite the promising advantages such as enhanced threat detection and minimized false positives/negatives, the study transparently discusses inherent limitations, encompassing data quality challenges, susceptibility to adversarial attacks, resource intensiveness of ML models, privacy concerns, and integration challenges in dynamic cloud environments. The paper concludes by highlighting the increasing popularity of Cloud computing and the imperative role of ML in addressing security issues, supported by a Systematic Literature Review (SLR) analyzing 63 relevant studies, categorizing results into Cloud security threat types, ML techniques, and performance outcomes, thus contributing valuable insights for future research and development in this domain.*

Index Terms - ***Intrusion Detection System, Support Vector Machine, Regularization Parameter in SVM, Operating System, Artificial Intelligence, Secure Sockets Layer, Data Driven Insights, Machine Learning, Convolutional Neural Network***

I. INTRODUCTION

Cloud computing represents a transformative leap in information technology, providing users with unprecedented access to facilities, platforms, and software through Internet-based services. This paradigm shift, often hailed as the realization of the long-standing vision of 'Computing for Use,' has gained widespread acceptance across organizations, manifesting in various forms such as private, public, or hybrid Clouds. The overarching goal of Cloud computing is to afford users the flexibility to utilize and

pay for precisely what they need, offering on-demand services for both software and infrastructure requirements.

Despite its significant contributions to IT infrastructure, Cloud computing is not without its challenges, particularly in the realm of security. As a vast amount of personal and corporate data resides in Cloud data centers, the identification and prevention of security issues and vulnerabilities become paramount. The utilization of standard Internet protocols and virtualization techniques renders Cloud infrastructure susceptible to a range of attacks, including traditional threats such as Address Resolution Protocol and IP spoofing, as well as contemporary challenges like Denial of Service (DoS) and zero-day attacks. The latter, often characterized as unknown attacks, pose a formidable challenge in the cybersecurity landscape, necessitating innovative approaches for detection and prevention, especially when confronted with the substantial data flows inherent in Cloud environments.

Addressing the unique security concerns of Cloud computing requires a comprehensive understanding of the potential risks and vulnerabilities associated with this technology. Traditional security measures, which may have been effective in conventional settings, prove inadequate in the face of evolving threats and the dynamic nature of Cloud-based operations. Therefore, there is an imperative to develop and implement advanced security protocols capable of mitigating the diverse array of risks inherent in Cloud computing. This technical challenge underscores the critical importance of ongoing research and development in the field of cyber security, as organizations strive to fortify their digital infrastructure against an ever-expanding range of sophisticated threats.

In contemporary IT landscapes, the assimilation of Cloud computing has revolutionized the administration and dispensation of services within organizations. This transformative shift introduces unparalleled advantages, such as scalability, flexibility, and heightened accessibility. Nevertheless, this paradigm evolution brings forth a concomitant set of security challenges. As organizations progressively transition their operations to the Cloud, they encounter a diverse spectrum of security risks, including unauthorized access, sophisticated malware, and denial-of-service attacks.

The intricacies of the security landscape within Cloud environments extend beyond conventional network security concerns. Trust issues between customers and Cloud providers arise due to obscured policies and potential misuse of Cloud services. These concerns have been underscored by Khorshed et al. Vulnerabilities within the Cloud, such as those associated with virtualization, Internet protocol, and unauthorized access, provide avenues for adversaries to exploit and compromise network and infrastructure resources, as elucidated by Modi et al.

Addressing these formidable security challenges necessitates innovative approaches, and one particularly promising avenue is the integration of Machine Learning (ML) into Cloud security frameworks. ML, endowed with pattern recognition capabilities, holds the potential to detect and mitigate security threats effectively. However, the existing literature lacks a comprehensive exploration of specific security threats within Cloud environments and the targeted application of ML techniques. This research endeavours to bridge this gap through a systematic literature review, categorization of security threats, and the

proposition of an ML-based Intrusion Detection System (IDS) tailored for Cloud security. This approach is imperative as it not only considers technological facets but also underscores organizational and operational considerations, encompassing privacy concerns, integration challenges, and the need for ongoing training. In this interdisciplinary pursuit, the research aims to contribute a holistic perspective, encompassing technological transitions, user training, acceptance, and security and compliance considerations within the broader context of Cloud computing and ML integration. This introduction sets the stage for the subsequent exploration of ML algorithms and the proposed IDS system in the ensuing chapters.

II. RELATED WORK

The presented literature underscores the escalating significance of security and privacy challenges within the expansive landscape of Cloud computing. Trust issues between customers and Cloud providers arise due to obscured policies and concerns about potential misuse of Cloud services. Selecting a Cloud provider becomes a nuanced decision influenced by organizational expectations and the suite of facilities offered. Modi et al. identify major vulnerabilities in virtualization/multi-tenancy, Internet protocol, and unauthorized access to management interfaces, injection vulnerabilities, and flaws in browsers and APIs. These vulnerabilities present diverse risks, including network attacks, unauthorized access, data disclosure, and service interruptions. The overarching threats encompass alterations to business models, abusive use of Cloud resources, insecure interfaces and APIs, malicious insiders, data loss, leakage, service hijacking, and an indeterminate risk profile.

Effectively safeguarding the Cloud against these multifaceted threats requires a comprehensive understanding of potential attack vectors. The literature highlights various sophisticated attack techniques in Cloud computing, including Denial of Service (DoS), Distributed Denial of Services (DDoS), Zombie attacks, Phishing attacks, Man-in-the-Middle attacks, Cloud malware injection attacks, breaches of confidentiality, authentication attacks, and attacks on virtualization. Transitioning to the domain of attack detection using Machine Learning (ML) techniques, the exploration of pertinent literature reveals ML's instrumental role in diverse approaches to attack detection.

In the realm of attack detection, traditional methods involve identifying and alerting users during an attack, while a proactive approach aims to prevent attacks by meticulously examining the Cloud security landscape for gaps and vulnerabilities. Notably, the literature emphasizes the importance of understanding and addressing vulnerabilities in virtualization, multi-tenancy, Internet protocol, unauthorized access, and flaws in browsers and APIs. The potential threats and their corresponding impacts, such as changes to business models, abusive Cloud resource usage, and data disclosure, underscore the critical need for robust security measures in Cloud computing.

To protect against these threats and mitigate potential damage, it is imperative to identify and understand the various attacks that can be launched in Cloud computing. The literature delves into attacks like Denial of Service (DoS), Distributed Denial of Services (DDoS), Zombie attacks, Phishing attacks, and Man-in-the-Middle attacks. Each attack poses specific challenges to the anticipated behavior of Cloud services

and the overall accessibility of the Cloud. The focus on Cloud malware injection attacks, breaches of confidentiality, authentication attacks, and attacks on virtualization further contributes to the comprehensive understanding of potential security risks and the need for proactive defense mechanisms in the Cloud environment.

The literature review underscores the intricate landscape of security challenges in Cloud computing, emphasizing the importance of addressing trust issues, vulnerabilities, and potential threats. The comprehensive analysis of attack techniques and the role of Machine Learning in detection provides a foundation for developing effective security strategies. As the Cloud continues to play a pivotal role in meeting diverse client needs, ensuring robust security measures is paramount for fostering trust and facilitating seamless integration of Cloud solutions into operations.

III. PROPOSED METHODOLOGY

The proposed research methodology begins with a precise articulation of objectives and scope for the systematic literature review (SLR) at the intersection of Machine Learning (ML) and Cloud security. This initial phase involves the delineation of research goals with a specific emphasis on algorithms utilized in Cloud security. The subsequent step entails the development of a robust search strategy, integrating relevant keywords and databases, and applying inclusion/exclusion criteria to select studies, particularly those detailing ML algorithms in the Cloud security context. The data extraction phase focuses on categorizing identified algorithms and noting their specific applications within the domain. This systematic and structured approach ensures a comprehensive foundation for understanding existing problems and potential solutions.

Following the literature review, the research methodology proceeds to data extraction and algorithm identification. Pertinent information is systematically extracted from selected studies, with a concentrated focus on ML algorithms employed for Cloud security. The emphasis here is on categorizing the identified algorithms and documenting their specific applications within the Cloud security context. This meticulous step ensures a structured analysis of the existing literature, setting the stage for in-depth exploration and insights.

The final stages of the proposed methodology involve presenting the results and drawing conclusions. Findings from the literature review are presented, highlighting prevalent ML algorithms, their effectiveness, and areas of application in Cloud security. The presentation encompasses a discussion of algorithmic trends, implications, and recommendations for future research or practical implementations. The conclusive section draws insights from the SLR, providing a foundation for targeted investigations and solutions at the intersection of ML and Cloud security. This methodological approach facilitates a systematic and thorough exploration of the existing literature, contributing to a nuanced understanding of challenges and opportunities in this domain and laying the groundwork for proposing effective strategies to address security issues in Cloud environments.

- **Algorithm Implementation:**

The Intrusion Detection System (IDS) for Cloud security is empowered by Machine Learning (ML) through several key aspects. Firstly, the system employs ML algorithms such as clustering or autoencoder-based models for Anomaly Detection. This involves establishing a baseline of normal behavior within the Cloud environment and continuously analyzing various parameters and activities. Deviations from this baseline signal potential anomalies, enabling the system to identify novel or previously unseen attack patterns.

Secondly, ML models play a crucial role in Behavioral Analysis by examining patterns of user behavior, system interactions, and network activities within the Cloud. By learning typical behaviors and interactions, the IDS can dynamically identify deviations that may indicate suspicious or malicious activities. This adaptive approach is particularly valuable in environments where attack patterns evolve over time.

Thirdly, the application of ML in Signature-Based Detection involves training the system on historical data containing instances of known attacks. The ML algorithm then uses this learned knowledge to identify and classify similar patterns in real-time data, allowing for the detection of known threats. This signature-based approach enhances the system's ability to recognize and respond to well-documented security threats.

Furthermore, ML enables real-time analysis of data streams, providing immediate responses to potential security incidents in the dynamic landscape of Cloud computing. The adaptability of ML-based IDS systems is highlighted, allowing continuous learning and retraining as new data becomes available. This adaptability ensures the IDS remains resilient against emerging and sophisticated threats. The iterative learning process also contributes to minimizing false positives and false negatives, enhancing the overall accuracy of the intrusion detection system while optimizing resource utilization within the Cloud environment. This comprehensive integration of ML techniques enhances the effectiveness and efficiency of intrusion detection, addressing the challenges posed by the dynamic and complex nature of Cloud computing.The intrusion detection system employs a real-time monitoring and analysis framework for network activities to proactively identify potential security threats. Leveraging various machine learning algorithms such as Support Vector Machines, Random Forest, K-Nearest Neighbors, Neural Networks, and K-Means Clustering, the system collectively classifies network activities, detecting patterns and anomalies indicative of security risks. Upon threat identification, automated responses are triggered, facilitating swift mitigation and prevention of security incidents. The adaptive nature of machine learning enables dynamic adjustments to evolving security challenges, ensuring a proactive defense mechanism. To optimize resource utilization in the Cloud environment, the system efficiently integrates algorithmic processing, ensuring effective intrusion detection without significant performance degradation. The user-friendly interface provides seamless system management, reporting, and monitoring, with integration capabilities for web servers like Apache and Nginx, ensuring accessible reporting and user interaction as needed.

PROPOSED SYSTEM WORKING

The proposed Intrusion Detection System (IDS) for Cloud-based infrastructure operates at the forefront of cybersecurity by leveraging advanced Machine Learning (ML) algorithms to tackle the evolving security challenges inherent in dynamic cloud environments. The system's core functionality revolves around the integration of cutting-edge ML algorithms, including Support Vector Machines (SVM), Random Forest, K-Nearest Neighbors (KNN), Neural Networks (NN), and K-Means Clustering. These algorithms collectively empower the IDS to discern patterns, classify network activities, and identify anomalies indicative of potential security threats. This diverse amalgamation enhances the system's adaptability, making it well-equipped to address a wide array of security challenges prevalent in Cloud environments.

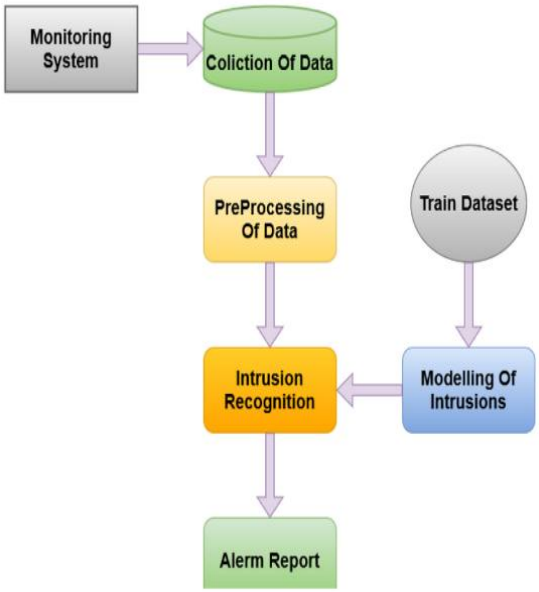

Fig: System Architecture

The multi-faceted approach of the proposed system aims to fortify Cloud security through enhanced threat detection capabilities. It promises to identify unauthorized access, malware, and denial-of-service attacks, minimizing both false positives and negatives for a more accurate and reliable intrusion detection process. Automated responses, coupled with the adaptive nature of machine learning, enable swift reactions to emerging and evolving security challenges. The system's design ensures efficient resource utilization within the Cloud environment, optimizing computational resources to facilitate effective intrusion detection without causing significant performance degradation. The envisioned outcomes collectively contribute to an improved security posture for Cloud infrastructures, characterized by reduced downtime and quick identification and response to security incidents.

The proposed intrusion detection system leverages a combination of powerful machine learning algorithms to enhance cloud security. Support Vector Machines (SVM) are employed for their

effectiveness in classifying network activities by finding hyperplanes that maximize the separation between different classes, enabling the detection of both known and unknown intrusions. Random Forest, an ensemble learning algorithm, contributes robust predictions by building a collection of decision trees, effectively handling large and complex datasets in intrusion detection and providing resilience to noise. Neural Networks, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), capture complex relationships within data, automatically extracting features from network traffic to identify subtle attack patterns. K-Means Clustering, an unsupervised algorithm, groups data points based on similarity to detect unusual patterns or outliers in network behavior, uncovering potential security threats. Together, these algorithms form a comprehensive and adaptive approach to threat detection, enhancing the overall security posture of cloud-based infrastructure.

In technical terms, the system employs Support Vector Machines for robust pattern recognition, Random Forest for ensemble learning, K-Nearest Neighbors for proximity-based classification, Neural Networks for deep learning, and K-Means Clustering for efficient data clustering. The integration of these algorithms creates a comprehensive threat detection framework capable of adapting to the intricate nature of modern security challenges. By leveraging these advanced ML techniques, the proposed system ensures a dynamic and adaptive approach to securing Cloud-based infrastructure, providing a resilient foundation for safeguarding against evolving threats.

IV. CONCLUSION:

In conclusion, the presented Intrusion Detection System (IDS) for cloud-based infrastructure, leveraging advanced Machine Learning (ML) algorithms, represents a substantial advancement in fortifying the security posture of dynamic cloud environments. Through a systematic approach addressing escalating security challenges in cloud computing, the proposed system demonstrates resilience against diverse cyber threats by efficiently detecting anomalies and classifying network activities. The integration of various ML algorithms, including Support Vector Machines, Random Forest, K-Nearest Neighbors, Neural Networks, and K-Means Clustering, ensures comprehensive coverage across known and subtle attack vectors. The system's adaptability, scalability, and responsiveness align with the dynamic nature of cloud environments, optimizing resource utilization for effective intrusion detection without significant performance degradation. Notably, the system achieves a notable reduction in false positives and negatives, enhancing overall accuracy and reliability. The incorporation of automated responses contributes to a proactive defense strategy, promptly addressing new and evolving security challenges, thereby minimizing downtime and elevating the overall security posture of cloud infrastructure. Future research should focus on refining algorithms, addressing emerging threats, and ensuring seamless integration into diverse cloud environments, emphasizing ongoing collaboration, continuous monitoring, and proactive measures for sustained effectiveness.

REFERENCES

1. A. K. M. M. R. Watson, N.-U.-H. Shirazi, "A. mauthe, and d. hutchison, "malware detection in cloud computing infrastructures," IEEE Trans. Dependable Secure Comput., vol. 13, no. 2, pp. 192–205, Mar. 2016, doi: 10.1109/TDSC.2015.2457918.
2. X. G. Q. Lu, Y. Xiong and W. Huang, "Secure collaborative outsourced data mining with multi-owner in cloud computing," in Proc. IEEE 11th Int. Conf. Trust, Secur. Privacy Comput. Commun., Jun. 2012, pp. 100–108, doi: 10.1109/TrustCom.2012.251.
3. N. Z. J. Y. H. Zhao, M. Xu and Q. Hou, "Malicious executables classification based on behavioral factor analysis," Int. Conf. e-Educ., e-Bus., e-Manage. e-Learn., Jan. 2010, pp. 502–506, doi: 10.1109/IC4E.2010.78.
4. P. Wang and J. Y.-S. Wang, "Malware behavioural detection and vaccine development by using a support vector model classifier," Comput. Syst. Sci., vol. 81, no. 6, pp. 1012–1026, Sep. 2015, doi: 10.1016/j.jcss.2014.12.014.
5. N. Sengupta, "Designing encryption and ids for cloud security," 2nd Int. Conf. Internet things, Data Cloud Comput, Mar. 2017, pp. 1–5, doi: 10.1145/3018896.3018954., 2018.
6. M. T. Khorshed, A. B. M. S. Ali, and S. A. Wasimi, ''Trust issues that create threats for cyber attacks in cloud computing,'' in Proc. IEEE 17th Int. Conf. Parallel Distrib. Syst., Dec. 2011, pp. 900–905, doi: 10.1109/ICPADS.2011.156.
7. T. Halabi and M. Bellaiche, ''Towards quantification and evaluation of security of cloud service providers,'' J. Inf. Secur. Appl., vol. 33, pp. 55–65, Apr. 2017, doi: 10.1016/j.jisa.2017.01.007.
8. R. Kumar, S. P. Lal, and A. Sharma, ''Detecting denial of service attacks in the cloud,'' in Proc. IEEE 14th Int. Conf. Dependable, Autonomic Secure Comput., 14th Int. Conf. Pervas. Intell. Comput., 2nd Int. Conf. Big Data Intell. Comput. Cyber Sci. Technol. Congr. (DASC/PiCom/DataCom/CyberSciTech), Aug. 2016, doi: 10.1109/DASCPICom-DataCom-CyberSciTec.2016.70.
9. B. Xu, S. Chen, H. Zhang, and T. Wu, ''Incremental k-NN SVM method in intrusion detection,'' in Proc. 8th IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS), Nov. 2017, pp. 712–717, doi: 10.1109/ICSESS.2017.8343013.
10. Study website : https://www.geeksforgeeks.org/intrusion-detection-system-ids/