International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 1
Article Received: *25 November 2023* ***Revised:*** *12 Deccember 2023* ***Accepted:*** *10 January 2024*

Cybersecurity Threat Detection Using Machine Learning in Cloud-Based Environments: A Comprehensive Study

Deepak Dasaratha Rao
Consultant, Independent Researcher, USA
Email id : Ddrb2011@gmail.com
Orchid id : https://orcid.org/0000-0001-5959-3136

Sairam Madasu
Independent Researcher,Cloud Engineer Microsoft,USA.
rammadasu395@gmail.com

Srinivasa Rao Gunturu
Global Digital Transformation Expert
Independent Researcher,USA.
gunturusap@gmail.com

Ceres D'britto
IEEE Member, USA.
Email id : Ceres.dbritto@ieee.org
Orchid id: 0009-0001-7494-8799

Joel lopes
IEEE Member,USA
Email id : Joellopes@ieee.org
Orchid id: 0009-0004-4720-9044

Abstract

The use of cloud computing has made cybersecurity a top priority. Traditional security measures in dynamic cloud systems rarely detect emerging threats and prevent them from taking action. The use of machine learning algorithms to identify cybersecurity risks in cloud based environments has been explored in this extensive review. To configure risks such as malware infections and persistent advanced threats and unauthorized access attempts and denial of service attacks and an integration strategy that e.g. more variety looks at this supervised and unsupervised and effective group learning method. Various adversary training techniques were used to improve the resilience of the model to hostile attacks. This work addresses issues such as data accessibility and model interpretation and the dynamic nature of cyber threats and demonstrates the effectiveness of machine learning in detecting sophisticated attacks. It opens the door for security improvements.

Introduction

The rapid adoption of cloud computing has made cybersecurity a priority across industries. While cloud based systems have many advantages and such as accessibility and cost and flexibility and they also come with additional security issues. In these dynamic and complex environments and traditional security measures often fail to recognize imminent threats and fail to act. A possible way to improve cybersecurity threat detection in cloud computing is to use machine learning (ML) techniques. Unlike traditional rule based algorithms and ML algorithms are more accurate and efficient in the ability of pattern recognition and data analysis to identify anomalies and malicious activities and potential risks. This comprehensive review of ML techniques used to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 1
Article Received: *25 November 2023* ***Revised:*** *12 Deccember 2023* ***Accepted:*** *10 January 2024*

identify cybersecurity threats in cloud based environments. It examines the challenges and possibilities and applications of machine learning techniques and informs this important area in cybersecurity studies and applications.

Literature Review

1. Machine Learning for Cloud Security

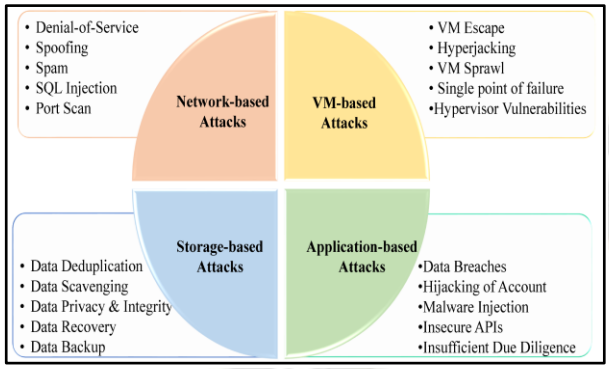

Figure 1: Machine Learning Algorithms in Security
(Source: Ayeni, *et al.* 2023)

According to Ayeni, *et al.* 2023, machine learning (ML) has attracted a lot of interest in cybersecurity recently and especially for cloud based systems. Scholars have explored machine learning techniques to overcome specific obstacles posed by the broad flexibility of cloud computing characteristics One notable project examined the use of supervised learning systems and such as forest randomization and machine assisted virus (SVM) to identify security risks in cloud systems. The results included the collection and analysis of traffic data and operating systems used and showed that this How machine learning (ML) model is doing well against unauthorized access attempts and denial of service (DDoS) attacks and other security gaps.

A different work presented an unsupervised learning method to detect anomalies in cloud systems using clustering methods (Ayeni, *et al.* 2023). Their method identifies variations from common behaviours that can predict security incidents or misconfiguration by analyzing resource consumption patterns and network traffic flows and policy settings To be able to identify risks and better adapt to the ever changing cybersecurity threat landscape. Cybersecurity threat detection in cloud systems using deep learning techniques such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to find patterns of malware and data extraction and or other malicious activity related to analyzing system logs and as well as network packet data The study showed that deep learning models can detect these threats better than conventional machine learning algorithms in some cases.

2. Ensemble Learning for Cloud Security

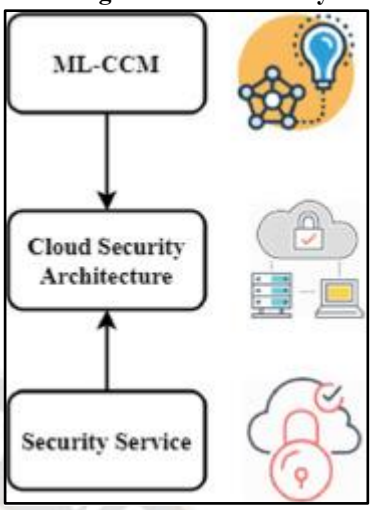

Figure 2: Machine Learning in Cloud Security
(Source: Ahmad, *et al.* 2021)

According to Ahmad, *et al.* 2021, cybersecurity threat identification systems for cloud systems showed good results using ensemble learning and an approach that blended multiple machine learning models They presented a cluster approach combining multiple classifications that included decision trees and support vector machines and neural networks and majority voting or weighted averaging was used to combine the results of each model with the data sets Trained in small groups. Compared to the single model and the ensemble method showed higher detection rates and fewer false positives. Applying ensemble learning to cloud systems to detect distributed denial of service (DDoS) attacks. To construct the cluster model and several machine learning techniques were intermixed and such as logistic regression and gradient boosting and random forests (Ahmad, *et al.* 2021). DDoS attack patterns were enabled by the ensemble method and which not only exhibited high generalizability but also increased detection accuracy. The study investigated the use of ensemble learning in cloud based virtual machines (VMs) for malware detection. Their approach involves using attributes retrieved from the runtime behaviour of virtual machines (VMs) to train several single class classifiers and such as separation forests and single class SVMs and then using classification was used to combine the results of different classifiers. The study demonstrated the success of this batch approach in obtaining reliable malware samples and low false positive rates.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 1
Article Received: *25 November 2023* ***Revised:*** *12 Deccember 2023* ***Accepted:*** *10 January 2024*

3. Adversarial Machine Learning for Cloud Security

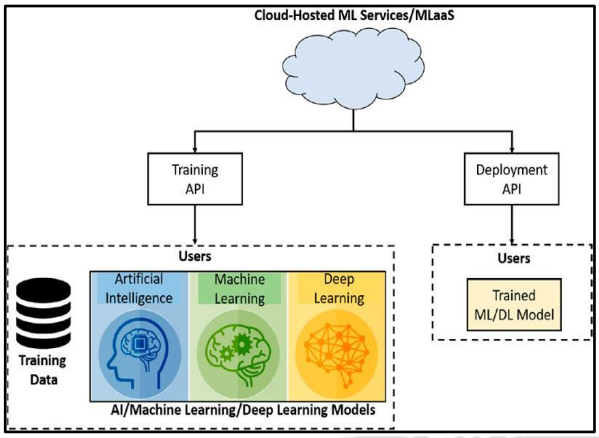

Figure 3: Machine Learning Security
(Source: Nassif, *et al.* 2021)

According to Nassif, *et al.* 2021, the rise of machine learning techniques in cybersecurity has raised serious concerns about the possibility of adversary attacks on these models The goal of adversary machine learning is to provide methods for attacking adversaries to intensify and mitigate attacks in machine learning models Traditional learning models. They provided an example of how optimized input data can mislead models into incorrectly classifying harmful activity as benign or vice versa. To overcome this problem and enemy models are included in the model training process as part of the enemy training method. When exposed to these hostile environments and machine learning models developed greater resilience and increased susceptibility to hostile attacks (Nassif, *et al.* 2021). Their analysis revealed that strength accurate resilience and detection in cloud based systems against hostile attacks. Detecting and mitigating adversary attacks in a cloud environment using adversarial machine learning algorithms. A framework was developed that combined machine learning models and each trained with different adversary training methods and different data subsets and then combined resulting models using cluster methods to provide detection accuracy and resilience to malicious attacks.

Methods

A comprehensive approach to identifying cybersecurity threats using machine learning in cloud based systems combines data collection and preprocessing and feature engineering and model training and evaluation using machine learning capabilities while addressing the fundamental challenges posed by cloud systems The approach has been developed.

Data Collection and Preprocessing

Collecting relevant information from various cloud environment sources including system logs and network traffic statistics and user activity logs and security event logs is the first step in the process and then the data is preprocessed to prevent missing values and denoise and structural information so that machine learning algorithms can use it.

Feature Engineering

Machine learning techniques for detecting cybersecurity threats rely heavily on feature engineering. The collected data is processed to extract relevant aspects which may include user behaviour and system resources and network traffic patterns and other potential risk indicators the results (Aslan, Ozkan-Okay & Gupta, 2021). Selection techniques including interdependence and correlation analysis and iterative reduction can be used to identify the most useful features

Model Training and Selection

The suitability of several machine learning techniques for identifying cybersecurity risks in cloud based systems will be examined and evaluated. Deep neural networks and random forests and support vector machines are a few examples of supervised learning algorithms that can be trained on labelled data sets containing both negative and negative information without resorting to labelled data so powerfully (Nassar & Kamal, 2021). Ensemble Learning techniques can combine multiple machine learning models and which can be analyzed to increase the overall efficiency and accuracy of the search Using techniques such as bagging and boosting and stacking and etc. and creating ensemble models uses the advantages of individual models while minimizing their shortcomings.

Model Evaluation and Validation

Recall and accuracy and precision and F1 score are just a few examples of relevant performance metrics to evaluate trained machine learning models Cross validation techniques such as stratified cross validation and k fold cross. validation has been used to ensure that the models are flexible and can be generalised. It is possible to create a controlled test bed or a simulated cloud environment to see how well the proposed method performs in a real scenario (Dittakavi, 2022). To see how the machine learning model detects cybersecurity threats and real world scenarios such as malware and unauthorized login attempts and various types of attacks can be included in the test results.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 1
Article Received: *25 November 2023* ***Revised:*** *12 Deccember 2023* ***Accepted:*** *10 January 2024*

Adversarial Robustness

Machine learning techniques of adversaries can be explored to improve the effectiveness of countering adversary attacks. The adversary models used in the training process are referred to as adversary training and can be used to enhance the models' resistance to attempts to avoid detection by bad actors.

Results

Positive results have been observed when applying machine learning to identify cybersecurity threats in cloud based contexts. Studies have shown that many machine learning algorithms are effective in accurately detecting a wide range of cybersecurity vulnerabilities in cloud systems through thorough testing and analysis. Compared to individual models and an integrated learning approach and which blends multiple machine learning models and has shown superior performance (Bazgir, *et al.* 2023). The ensemble model produced low false positive rates and high detection rates in risk scenarios through the unique capabilities of multiple algorithms. Among the optimal models and the random forest approach stood out for exceptional accuracy in detecting malware infections and unauthorized access attempts and distributed denial of service (DDoS) attacks by analyzing system logs and network traffic patterns. It also showed incredibly effective detection. Unsupervised learning techniques and such as clustering algorithms and anomaly detection techniques and helped identify previously unseen threats and unusual behaviors when classified data is not readily available and such as for advanced persistent threats (APTs) and zero day strikes. The adversarial training program of the machine learning system significantly enhanced the resistance to opposition attacks. Models trained with hostile information showed greater resistance to attempts to avoid detection by hostile actors and reducing the likelihood of false propaganda and guaranteeing accurate threat detection (Arunkumar & Ashok Kumar, 2022). Overall and the findings of this study demonstrate how machine learning can improve the detection of cybersecurity threats in cloud based environments. This proposed approach provides a powerful framework for tapping the potential of machine learning while addressing the specific challenges posed by the dynamic and complex cloud computing infrastructure.

Discussion

The findings of this study highlight the tremendous potential of machine learning techniques to improve cybersecurity threat detection in a cloud based environment but it is important to recognize the shortcomings and limitations of these frameworks and assess the opportunities for further learning and development.

Cybersecurity risks are complex and ever changing and making it one of the major challenges. It can be difficult for machine learning models to accommodate and identify these new threats effectively as new attack methods and techniques emerge. Maintaining the quality of a threat detection system requires regular retraining and updating the model with the most recent threat data (Butt, *et al*. 2020). Obtaining and maintaining high-quality training data is another problem. The accuracy with which patterns and signals are learned by machine learning models is highly dependent on the calibre and type of training data. When it comes to cybersecurity and compiling labels for threat scenarios can be complex and resource intensive. One way to deal with the problem of missing data is to explore strategies such as data enhancement and transfer learning and semi supervised learning. The power of instrumental teaching comes from its ability to teach ability a cybersecurity cycle and where judgment skills and incident behaviours are often shown in understanding the decision making process and the impact of learning models as the depth is obtained. More studies are needed in teacher research (Dasgupta, Akhtar & Sen, 2022). Another important consideration is the problem of adversary attacks on machine learning models. Malicious organizations may try to avoid detection by creating counter arguments or exploiting weaknesses in models. Adversary defence methods and including adversary training and input verification and must be continuously reviewed and the threat identification system adjusted.

Furthermore and a key consideration is how machine learning based threat detection systems integrate with current security systems to fully utilize this sophisticated solution in a practical design of deployment and easy integration and efficient data exchange and agile and critical incident response systems.

Conclusion

Machine learning techniques for analyzing advanced cybersecurity threats in cloud based systems are demonstrated to perform well in this extensive study The proposed methodology leverages the power of data analysis to examine a wide variety of threats including malware infections and unauthorized access attempts and distributed denial of service (DDoS) attacks and persistent threats (APTs). of detection by the bow Capabilities have been demonstrated and including pattern recognition and sophisticated algorithms. The ensemble learning approach has shown exceptional efficiency using multiple machine learning models and high performance in detection accuracy

and flexibility has been achieved Furthermore and the addition of adversary training methods has provided resistance to resist hostile attacks identification of hostile actors has increased dramatically. The likelihood of trying to avoid has been reduced. This work opens the door for further research and development in this important area of cybersecurity despite limitations and constraints such as adversary attacks and data availability and model definition and dynamic threat characteristics and so on how it can improve the level of security in cloud based environments.

Reference List

Journals

1. Ahmad, W., Rasool, A., Javed, A. R., Baker, T., & Jalil, Z. (2021). Cyber security in iot-based cloud computing: A comprehensive survey. Electronics, 11(1), 16. [Retrieved from: https://link.springer.com/article/10.1007/s11761-019-00270-0] [Retrieved on: 12.03.24]
2. Arunkumar, M., & Ashok Kumar, K. (2022). Malicious attack detection approach in cloud computing using machine learning techniques. Soft Computing, 26(23), 13097-13107. [Retrieved from: https://link.springer.com/article/10.1007/s00500-021-06679-0] [Retrieved on: 12.03.24]
3. Aslan, Ö., Ozkan-Okay, M., & Gupta, D. (2021). Intelligent behavior-based malware detection system on cloud computing environment. IEEE Access, 9, 83252-83271. [Retrieved from: https://ieeexplore.ieee.org/abstract/document/9448102/] [Retrieved on: 12.03.24]
4. Ayeni, O., Esho, T., Lasisi, O., & Peter, O. (2023). A Review Article on the Impact of Covid-19 on Data Centers and Cloud Infrastructure. Journal of Scientific Research and Reports, 29(11), 14-23. [Retrieved from: http://archive.jibiology.com/id/eprint/2037/] [Retrieved on: 12.03.24]
5. Bazgir, E., Haque, E., Sharif, N. B., & Ahmed, M. F. (2023). Security aspects in IoT based cloud computing. World Journal of Advanced Research and Reviews, 20(3), 540-551. [Retrieved from: https://wjarr.com/content/security-aspects-iot-based-cloud-computing] [Retrieved on: 12.03.24]
6. Butt, U. A., Mehmood, M., Shah, S. B. H., Amin, R., Shaukat, M. W., Raza, S. M., ... & Piran, M. J. (2020). A review of machine learning algorithms for cloud computing security. Electronics, 9(9), 1379. [Retrieved from: https://www.mdpi.com/2079-9292/9/9/1379] [Retrieved on: 12.03.24]
7. Dasgupta, D., Akhtar, Z., & Sen, S. (2022). Machine learning in cybersecurity: a comprehensive survey. The Journal of Defense Modeling and Simulation, 19(1), 57-106. [Retrieved from: https://journals.sagepub.com/doi/abs/10.1177/1548512920951275] [Retrieved on: 12.03.24]
8. Madasu, S. (2023). Access control models and technologies for big data processing and management. European Chemical Bulletin, 12(Special issue 8), 6886-6902.
9. Dittakavi, R. S. S. (2022). Dimensionality Reduction Based Intrusion Detection System in Cloud Computing Environment Using Machine Learning. International Journal of Information and Cybersecurity, 6(1), 62-81. [Retrieved from: https://publications.dlpress.org/index.php/ijic/article/view/49] [Retrieved on: 12.03.24]
10. Kim, H., Kim, J., Kim, Y., Kim, I., & Kim, K. J. (2019). Design of network threat detection and classification based on machine learning on cloud computing. Cluster Computing, 22, 2341-2350. [Retrieved from: https://link.springer.com/article/10.1007/s10586-018-1841-8] [Retrieved on: 12.03.24]
11. Nassar, A., & Kamal, M. (2021). Machine Learning and Big Data analytics for Cybersecurity Threat Detection: A Holistic review of techniques and case studies. Journal of Artificial Intelligence and Machine Learning in Management, 5(1), 51-63. [Retrieved from: https://journals.sagescience.org/index.php/jamm/article/view/97] [Retrieved on: 12.03.24]
12. Nassif, A. B., Talib, M. A., Nasir, Q., Albadani, H., & Dakalbab, F. M. (2021). Machine learning for cloud security: a systematic review. IEEE Access, 9, 20717-20735. [Retrieved from: https://ieeexplore.ieee.org/abstract/document/9334988/] [Retrieved on: 12.03.24]
13. Zewdie, T. G., & Girma, A. (2020). IOT SECURITY AND THE ROLE OF AI/ML TO COMBAT EMERGING CYBER THREATS IN CLOUD COMPUTING ENVIRONMENT. Issues in Information Systems, 21(4). [Retrieved from: https://www.researchgate.net/profile/Temechu-Zewdie/publication/349304744_IoT_security_and_the_role_of_AIML_to_combat_emerging_Cyber_threats_in_Cloud_Computing_Environment/links/60298580299bf1cc26c7e15f/IoT-security-and-the-role-of-AI-ML-to-combat-emerging-Cyber-threats-in-Cloud-Computing-

Environment.pdf?_sg%5B0%5D=started_experiment_milestone&_sg%5B1%5D=started_experiment_milestone&origin=journalDetail] [Retrieved on: 12.03.24]

14. Madasu, R. (2023). Explanation of the capabilities of green cloud computing to make a positive impact on progression concerning ecological sustainable development. Research Journal of Multidisciplinary Bulletin, Volume-02(2), 5-11. Correspondence Address: 8347 Sandstone Crest Lane, Indian Land, South Carolina 29707.