IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

Review Paper on Cloud Intrusion Detection System

Prof. R. G. Waghmare[1], Kaustubh M. Karale[2], Omkar A. Raut[3]
Professor, Department of Artificial Intelligence and Machine Learning[1]
Students, Department of Artificial Intelligence and Machine Learning[2,3]
All India Shri Shivaji Memorial Society Polytechnic Pune, Maharashtra, India

Abstract: *The study proposes an enhanced cloud intrusion detection system (IDS) that tackles security challenges in cloud computing, focusing on data imbalance and feature selection. By integrating SMOTE for data imbalance and a hybrid feature selection method, the system achieves exceptional accuracies exceeding 98% and 99% on two datasets. The use of fewer informative features enhances system efficiency, showcasing its practical applicability and effectiveness in real-world scenarios. Overall, the study contributes significantly to cloud security by offering a holistic approach to IDS enhancement.*

Keywords: cloud computing, intrusion detection system, security, data imbalance, feature selection, SMOTE, hybrid method, accuracy, efficiency, real-world scenarios, cloud security

I. INTRODUCTION

In recent years, the landscape of digital technology has been reshaped by the advent of cloud computing (CC). This Review delves into the intricate relationship between CC and cybersecurity, offering a comprehensive analysis of the evolving threat landscape and the measures taken to combat it.

The review begins by highlighting the expansive reach of CC across various industries, attributed to its versatility in offering services such as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). The benefits of CC, including scalability and cost-effectiveness, are duly acknowledged, setting the stage for a deeper exploration of its security implications.

One of the standout features of this review is its meticulous examination of the escalating cybers ecurity threats accompanying the proliferation of CC. With research findings indicating a significant surge in cyberattacks over the pastdecade, the urgency to fortify cloud infrastructures against such threats becomes apparent. The review underscores t heproactive stance taken by cloud service providers (CSPs) in bolstering cybersecurity measures to safeguard user data.

A particularly insightful aspect of this review is its discussion on the economic ramifications of security breaches, emphasizing the substantial financial losses incurred by companies and organizations. By contextualizing the global cybersecurity spending and projected growth rates, the review underscores the magnitude of investments required to mitigate security risks in CC environments.

Furthermore, the review delves into the pivotal role played by networks in securing cloud environments, elucidating the diverse cybersecurity strategies employed to safeguard against potential attacks. Traditional security measures are scrutinized for their limitations in countering sophisticated threats, paving the way for the emergence of anomaly-based Intrusion Detection Systems (IDSs) powered by machine learning (ML) algorithms

The review offers a nuanced exploration of ML-based intrusion detection approaches, elucidating their self-learning capabilities and adaptability to evolving threats. By dissecting the performance of supervised ML techniques in identifying various types of attacks, the review underscores the importance of robust feature selection strategies to optimize modelperformance and mitigate imbalanced data issues.

A highlight of this review is its evaluation of a hybrid feature selection strategy using real -world datasets, showcasing the efficacy of Random Forest (RF) classifiers in detecting and classifying cyber threats. The promising results underscore thepotential o ML-based approaches in fortifying cloud infrastructures against cybersecurity threats.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

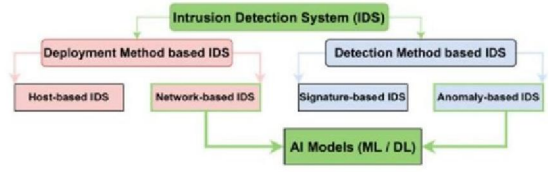

FIGURE 1. IDS classification.

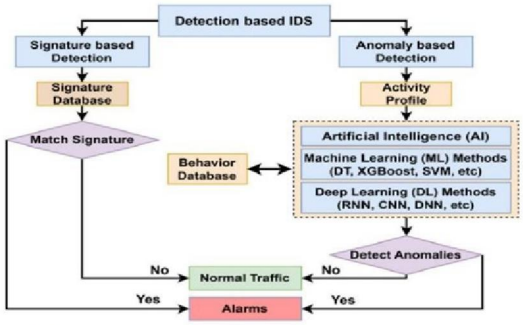

FIGURE 2. Taxonomy of IDS based on the detection method.

FIGURE 3. Process of building an ML model.

II. EXISTING SYSTEM

Recent studies have focused on improving Intrusion Detection Systems (IDS) performance through various feature selection techniques and machine learning (ML) and deep learning (DL) classifiers. Here's a brief overview of each study:

1. Benmessahel et al.: Introduced an evolutionary neural network (ENN) using a multiverse optimizer (MVO) algorithm, demonstrating high efficacy in recognizing new threats, especially on the UNSW-NB15 dataset.
2. Yang et al.: Presented a model incorporating a Deep Neural Network (DNN) with an Improved Conditional Variational Autoencoder (IC-VAE), showcasing significant performance improvements in detecting unknown and minority attacks.
3. Tama et al.: Proposed a hybrid IDS system utilizing feature selection techniques like Particle Swarm Optimization (PSO), Ant Colony Algorithm (ACA), and Genetic Algorithm (GA), coupled with a two-level meta-ensemble classifier for considerable performance enhancement.
4. Khan et al. : Introduced a two-stage deep learning approach (TSDL) employing a Stacked Auto-Encoder (SAE) and a Soft-max classifier, effectively addressing unbalanced data using the SMOTE algorithm.
5. Vinayakumar et al.: Suggested a hybrid IDS alert system utilizing distributed deep learning algorithms, outperforming traditional machine learning classifiers in most cases, with decision trees and random forests showing superior performance on the UNSW-NB15 dataset.
6. Patil et al. : Proposed a framework for hypervisor-level distributed network security (HLDNS) effectively detecting both known and unknown threats while minimizing computing costs.
7. Saleh et al. : Proposed an IDS methodology combining naïve base feature selection (NBFS), optimized support vector machines (OSVM), and prioritized k-nearest neighbors (PKNN) techniques for real-time threat detection and multi-class classification.
8. Zhang et al. : Suggested an MSCNN-LSTM methodology utilizing multi-scale convolutional neural networks and long short-term memory networks for spatial-temporal feature analysis, showcasing promising results on complex datasets.
9. Kasongo and Sun : Utilized the XGBoost model for feature selection and various ML methods for threat classification, highlighting the effectiveness of decision trees (DT) and artificial neural networks (ANN) for binary and multi-class classification.
10. Kumar et al. : Proposed an integrated classification-based IDS utilizing decision trees (DT) for classification and Information Gain (IG) for feature selection, demonstrating performance on both offline and online datasets.
11. Almomani : Designed an IDS based on bio-inspired feature selection techniques and machine learning classifiers, showcasing improved results through feature selection strategies, particularly using the Particl e Swarm Optimization (PSO) algorithm.
12. Jiang et al. : Discussed a technique integrating hybrid sampling and deep hierarchical networks for IDS, effectively balancing datasets and producing high-quality results by extracting spatial and temporal attributes.
13. Rajesh Kanna and Santhi : Suggested an optimized CNN-HMLSTM model for spatial-temporal threat detection, facing challenges in training due to complexity, prompting future investigations into feature selection procedures.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

14. Sreelatha et al. : Presented an efficient cloud IDS employing feature selection and classification techniques, demonstrating superior performance, with plans to enhance effectiveness through hybrid optimization algorithms.
15. Kanna and Santhi : Proposed an IDS based on hybrid-optimized deep learning involving artificial bee colony methods and convolutional long short-term memory networks, requiring significant training and testing times despite notable performance

III. IMPLEMENTATION

This section outlines a meticulous approach to preprocessing datasets for training an Intrusion Detection System (IDS) in the cloud. The authors recognize the significance of dataset preprocessing due to its large size and diverse content, which includes both attack and non-attack data.

They propose a multifaceted feature selection strategy combining filter-based (IG and CS) and bio-inspired-based (PSO) methods. This comprehensive approach aims to identify the most relevant features for training the classifier, crucial for accurately distinguishing benign packets from attacks.

The preprocessing steps include the removal of irrelevant features, handling null values, and encoding categorical features. Additionally, feature scaling is applied to normalize values, ensuring consistency in the dataset. To address data imbalance issues, the authors employ the Synthetic Minority Over-sampling Technique (SMOTE), enhancing the model's ability to handle skewed datadistributions.

The feature selection process involves Information Gain (IG), Chi Square (CS), and Practical Swarm Optimization (PSO) techniques, aiming to optimize feature subsets for model training. Finally, the Random Forest (RF) classifier is utilized on the preprocessed dataset to detect intrusions.

Despite the effectiveness demonstrated by the approach, the authors acknowledge several limitations, including the potential for overfitting, sensitivity to outliers, and the need for careful hyperparameter selection for the RF classifier. These limitations highlight areas for future improvement and refinement of the proposed methodology

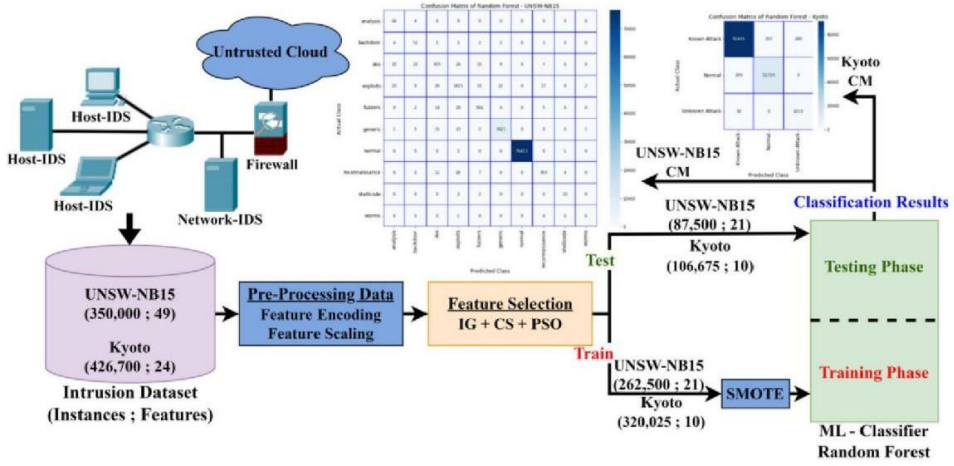

FIGURE 4. Overview of the suggested intrusion detection system based on a hybrid of feature selection methods.

IV. CONCLUSION

The focus on cloud security has led to leveraging machine learning for intrusion detection, distinguishing it from deep learning due to resource efficiency. Combining various feature selection algorithms enhances intrusion detection accuracy, as demonstrated by a proposed system. Future work aims to employ deep learning, ensemble learning, and meta-heuristic optimization for improved performance, tested on recent datasets reflecting contemporary network threats.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

REFERENCES

[1] R. R. Kumar, A. Tomar, M. Shameem, and M. N. Alam, ‘‘OPTCLOUD: An optimal cloud service selection framework using QoS correlation lens,’’ Comput. Intell. Neurosci., vol. 2022, pp. 1–16, May 2022, doi: 10.1155/2022/2019485.

[2] R. R. Kumar, M. Shameem, R. Khanam, and C. Kumar, ‘‘A hybrid evaluation framework for QoS based service selection and ranking in cloud environment,’’ in Proc. 15th IEEE India Council Int. Conf., Oct. 2018, pp. 1 –6, doi: 10.1109/INDICON45594.2018.8987192.

[3] M. Bakro, S. K. Bisoy, A. K. Patel, and M. A. Naal, ‘‘Performance analysis of cloud computing encryption algorithms,’’ in Advances in Intelligent Computing and Communication, in Lecture Notes in Networks and Systems, vol. 202. Singapore: Springer, 2021, pp. 357–367, doi: 10.1007/978-981-16-0695-3_35.

[4] (2020). Malware Statistics & Trends Report | AV-TEST. Accessed: Jan. 21, 2023. [Online]. Available: https://www.av-test.org/en/statistics/ malware/

[5] Digital Technology Market Research Services | Juniper Research. Accessed: Jan. 21, 2023. [Online]. Available: https://www.juniperresearch. com/home

[6] Cyber Security Market Size, Share & Trends Report, 2030. Accessed: Jan. 21, 2023. [Online]. Available: https://www.grandviewresearch. com/industry-analysis/cyber-security-market

[7] R. R. Kumar, M. Shameem, and C. Kumar, ‘‘A computational framework for ranking prediction of cloud services under fuzzy environment,’’ Enterprise Inf. Syst., vol. 16, no. 1, pp. 167–187, 10.1080/17517575.2021.1889037. Jan. 2022, doi:

[8] M. A. Akbar, M. Shameem, S. Mahmood, A. Alsanad, and A. Gumaei, ‘‘Prioritization based taxonomy of cloud - based outsource software development challenges: Fuzzy AHP analysis,’’ Appl. Soft Comput., vol. 95, Oct. 2020, Art. no. 106557, doi: 10.1016/j.asoc.2020.106557.

[9] M. Bakro, R. R. Kumar, A. A. Alabrah, Z. Ashraf, S. K. Bisoy, N. Parveen, S. Khawatmi, and A. Abdelsalam, ‘‘Efficient intrusion detection system in the cloud using fusion feature selection approaches and an ensemble classifier,’’ Electronics, vol. 12, no. 11, p. 2427, May 2023, doi: 10.3390/electronics12112427.

[10] M. Bakro, S. K. Bisoy, A. K. Patel, and M. A. Naal, ‘‘Hybrid blockchainenabled security in cloud storage infrastructure using ECC and AES algorithms,’’ in Blockchain based Internet of Things. Singapore: Springer, 2022, pp. 139–170, doi: 10.1007/978-981-16-9260-4_6.

[11] Z. Ahmad, A. S. Khan, C. W. Shiang, J. Abdullah, and F. Ahmad, ‘‘Network intrusion detection system: A systematic study of machine learning and deep learning approaches,’’ Trans. Emerg. Telecommun. Technol., vol. 32, no. 1, p. e4150, Jan. 2021, doi: 10.1002/ett.4150.

[12] I. F. Kilincer, F. Ertam, and A. Sengur, ‘‘Machine learning methods for cyber security intrusion detection: Datasets and comparative study,’’ Comput. Netw., vol. 188, Apr. 2021, Art. no. 107840, doi: 10.1016/j.comnet.2021.107840.

[13] I. Benmessahel, K. Xie, and M. Chellal, ‘‘A new evolutionary neural networks based on intrusion detection systems using multiverse optimization,’’ Int. J. Speech Technol., vol. 48, no. 8, pp. 2315 –2327, Aug. 2018, doi: 10.1007/S10489-017-1085-Y.

[14] Y. Yang, K. Zheng, C. Wu, and Y. Yang, ‘‘Improving the classification effectiveness of intrusion detection by using improved conditional variational AutoEncoder and deep neural network,’’ Sensors, vol. 19, no. 11, p. 2528, Jun. 2019, doi: 10.3390/s19112528.

[15] B. A. Tama, M. Comuzzi, and K. Rhee, ‘‘TSE-IDS: A twostage classifier ensemble for intelligent anomaly-based intrusion detection system,’’ IEEE Access, vol. 7, pp. 94497–94507, 2019, doi: 10.1109/ACCESS.2019.2928048.

[16] F. A. Khan, A. Gumaei, A. Derhab, and A. Hussain, ‘‘TSDL: A two-stage deep learning model for efficient network intrusion detection,’’ IEEE Access, vol. 7, pp. 30373–30385, 2019, doi: 10.1109/ACCESS.2019.2899721.

[17] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, ‘‘Deep learning approach for intelligent intrusion detection system,’’ IEEE Access, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[18] R. Patil, H. Dudeja, and C. Modi, ‘‘Designing an efficient security framework for detecting intrusions in virtual network of cloud computing,’’ Comput. Secur., vol. 85, pp. 402–422, Aug. 2019, doi: 10.1016/j.cose.2019.05.016.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 3, April 2024

[19] A. I. Saleh, F. M. Talaat, and L. M. Labib, ‘‘A hybrid intrusion detection system (HIDS) based on prioritized k - nearest neighbors and optimized SVM classifiers,’’ Artif. Intell. Rev., vol. 51, no. 3, pp. 403 –443, Mar. 2019, doi: 10.1007/s10462-017-9567-1.

[20] J. Zhang, Y. Ling, X. Fu, X. Yang, G. Xiong, and R. Zhang, ‘‘Model of the intrusion detection system based on the integration of spatial– temporal features,’’ Comput. Secur., vol. 89, Feb. 2020, Art. no. 101681, doi: 10.1016/j.cose.2019.101681.

[21] S. M. Kasongo and Y. Sun, ‘‘Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 dataset,’’ J. Big Data, vol. 7, no. 1, pp. 1–12, Dec. 2020, doi: 10.1186/s40537-020- 00379-6.

[22] V. Kumar, D. Sinha, A. K. Das, S. C. Pandey, and R. T. Goswami, ‘‘An integrated rule based intrusion detection system: Analysis on UNSWNB15 data set and the real time online dataset,’’ Cluster Comput., vol. 23, no. 2, pp. 1397 – 1418, Jun. 2020, doi: 10.1007/s10586-019-03008-x.

[23] O. Almomani, ‘‘A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms,’’ Symmetry, vol. 12, no. 6, pp. 1–20, 2020, doi: 10.3390/sym12061046.

[24] K. Jiang, W. Wang, A. Wang, and H. Wu, ‘‘Network intrusion detection combined hybrid sampling with deep hierarchical network,’’ IEEE Access, vol. 8, pp. 32464–32476, 2020, doi: 10.1109/ACCESS.2020.2973730.

[25] P. Rajesh Kanna and P. Santhi, ‘‘Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features,’’ Knowl.-Based Syst., vol. 226, Aug. 2021, Art. no. 107132, doi: 10.1016/j.knosys.2021.107132.

[26] G. Sreelatha, A. V. Babu, and D. Midhunchakkaravarthy, ‘‘Improved security in cloud using sandpiper and extended equilibrium deep transfer learning based intrusion detection,’’ Cluster Comput., vol. 25, no. 5, pp. 3129–3144, Oct. 2022, doi: 10.1007/s10586-021-03516-9.

[27] P. R. Kanna and P. Santhi, ‘‘Hybrid intrusion detection using MapReduce based black widow optimized convolutional long short-term memory neural networks,’’ Expert Syst. Appl., vol. 194, May 2022, Art. no. 116545, doi: 10.1016/j.eswa.2022.116545.

[28] S. Krishnaveni, S. Sivamohan, S. S. Sridhar, and S. Prabakaran, ‘‘Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing,’’ Cluster Comput., vol. 24, no. 3, pp. 1761–1779, Sep. 2021, doi: 10.1007/s10586-020-03222-y.

[29] K. Potdar, ‘‘A comparative study of categorical variable encoding techniques for neural network classifiers,’’ Int. J. Comput. Appl., vol. 175, no. 4, pp. 7–9, Oct. 2017, doi: 10.5120/ijca2017915495.

[30] M. Rashid, J. Kamruzzaman, T. Imam, S. Wibowo, and S. Gordon, ‘‘A tree-based stacking ensemble technique with feature selection for network intrusion detection,’’ Int. J. Speech Technol., vol. 52, no. 9, pp. 9768 –9781, Jul. 2022, doi: 10.1007/s10489-021-02968-1.