Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 3
Open Access

Adaptive Intrusion Detection Mechanisms For Enhancing Security In Cloud-Hosted Big Data Systems

Dr.J Jabez[1],Ahmed Alkhayyat[2], Dr.G.Vengatesan[3] Dr.R. Vasanthan[4]
[1].PROFESSOR, Department of Computer Science, Sathyabhama Institute of Science and Technology, Chennai. Jabezme@gmail.com
[2]. College of technical engineering, The Islamic university, Najaf, Iraq . ahmedalkhayyat85@iunajaf.edu.iq
[3].Associate Professor and Head, Department of Commerce (Business Analytics) KPR College of Arts Science and Research, Avinashi Road, Arasur, Coimbatore - 641407. dr.g.vengatesan@gmail.com
[4].Associate Professor, Department of EnglishNagaland University, Kohima Campus-797001.Nagaland, India.
Email:vasanthan@nagalanduniversity.ac.in. orcid: https://orcid.org/0000-0002-0962-3453

Cite this paper: Dr .J Jabez, Ahmed Alkhayyat , Dr. G.Vengatesan, Dr. R. Vasanthan (2024) Adaptive Intrusion Detection Mechanisms For Enhancing Security In Cloud-Hosted Big Data Systems. *Frontiers in Health Informatics, 13* (3), 5313-5327

Abstract
The development and evaluation of a new Intrusion Detection System (IDS) using CNN and SVM is the main contribution of this paper. The hybrid approach using the feature extraction strengths of CNNs for very powerful invariant features is proposed to identify complex patterns that indicate possible intrusions in achieving automated network traffic data analysis. The model enhances detection with higher effectiveness across different attack vectors by incorporating the use of SVM as a classifier. Extensive experiments have been carried out to compare the performance of the proposed CNN+SVM model against traditional IDS methodologies. The results indicate a significant improvement in the detection rate, accuracies reaching a high of 98% with this hybrid model while keeping false positives to a minimum, whereas true positive cases are maximized. Other measures, such as precision, recall, and F1-score, further establish the fact that a model can compete by striving to balance sensitivity and specificity in order to handle one of the most critical challenges that cybersecurity faces. These findings demonstrate the potential benefit of deep learning models coupled with traditional machine learning techniques for intrusion detection systems. The contribution of this work to the extension in the literature on IDS, and above all, to some important practical implications in securing network infrastructures from the new, continuously evolving cyber threats, is important. Future studies will focus on model optimization and its application in different network scenarios.

Keywords: *Intrusion Detection, Security, Cloud, Big Data, and Machine Learning.*

Introduction

Intrusion detection can be one of the most paramount things an information system can undertake in the identification of specific malicious activities that would compromise or damage information systems [1, 2]. Intrusion Detection Systems are deployed for system and network behavior monitoring in order to look out for suspicious activities that may signify an ongoing attack.

Generally, Intrusion Detection Systems have been classified into two broad categories: host-based IDS and

network-based IDS. On the other hand, the host-based IDS is interested in the events of a single computer or server from the inside-out, looking at its log files, system events, and application usage to notice unauthorized changes. In general, an IDS that is host-based monitors these internal aspects continuously and can therefore detect attacks or misuse that do not appear at the network level. This hosts a very powerful layer of security for the individual system.

Network-based IDS performs its detection in a more general approach: by analyzing network traffic to identify patterns or anomalies indicative of malicious activities across the entire network. Examples of network-based IDS include monitoring the flow of data across all devices on a network, looking for packet content that could indicate a well-known cyberattack [3]. Denial-of-Service (DoS) tries to overwhelm the system with traffic. SQL injection is an attempt to get into the database by submitting malicious material through input fields. Password attacks use brute force methods in an attempt to gain unauthorized access. Network-based IDS detects attacks in real time by inspecting network traffic, so immediate responses can be made before an attack is perpetrated inside or affects critical systems.

With a rapid increase in the reliance on interconnected networks and an exponential growth of network applications globally, it is believed that sophisticated cyberattacks are also on the rise. As more organizations embark on digital transformation, this increases the attack surface, hence making the likeliness of threats shift toward network vulnerabilities [4]. In that respect, an effective IDS must be either host-based or network-based and developed on a continuous basis in order to match the fast-changing profile of cybersecurity threats. This would mean strong protection not only at the individual system level but throughout whole networks from unauthorized access and attacks.

Intrusion detection has, in general, been involved with identifying unauthorized access, malicious attack, or misuse of an operating system or network to prevent security breach that can lead to either data compromise, disruption of services, or financial loss and damage to reputation. However, with evolving cyber threats at such high speeds, it is becoming tough to design an efficient and reliable IDS system, especially in the face of sophisticated attack techniques like distributed denial-of-service attacks, ransomware, and advanced persistent threats [5, 6]. These systems must grapple with the massive volumes of data coming off modern networks; they must find the faint attack patterns, perhaps explicitly crafted to evade traditional detection methods. It needs to be done in real time, adapting to novel and emerging threats. Combining this with the complexity of the imperative for accurate and timely detection underlines the problem of enhancing IDS capabilities in keeping with ever-growing sophistication of cyber-attacks.

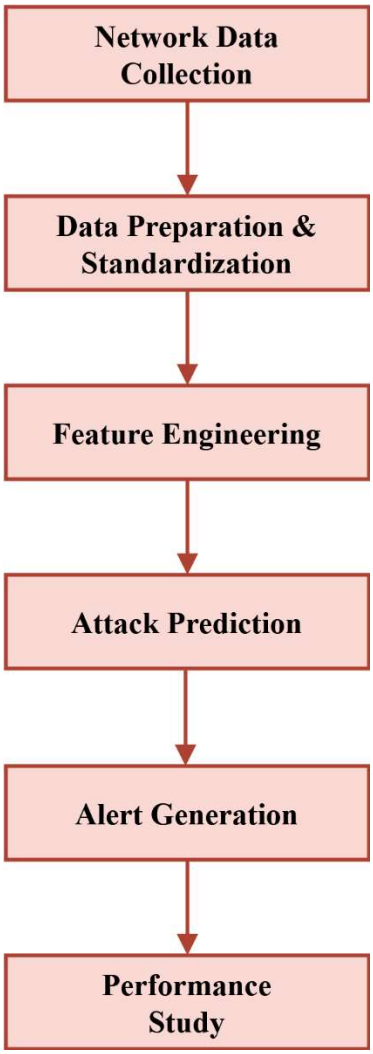

Figure 1. Typical attack detection process

This is further motivated by the fact that undetected or late detection of cyber-attacks has enormous consequences. The aftermath of a successful intrusion may come in the form of stolen sensitive data, system unavailability, loss of money, and breach of privacy [7]. The stakes are exceptionally high for an enterprise, government institution, or an individual since cyber criminals increasingly attack not only financial systems but also critical infrastructure, healthcare networks, and intellectual property. Cyberattacks vary and complicate further; hence, reliance on traditional security mechanisms such as firewalls or antivirus is not good enough.

There is an actual need for advanced intrusion detection methods that could actively identify and contain threats to guarantee information integrity, availability, and confidentiality. Furthermore, the ever-growing network traffic, which is driven by the increased usage of IoT devices, mobile applications, and cloud computing, increases the attack surface and, consequently, the need for intrusion detection systems. In order for them to actively protect such environments [8, 9], there is a need to develop IDS solutions that keep pace with network demands through their ability to make a correct separation between benign and malicious activities and allow for real-time threat detection and response. It therefore calls for further research and development in intrusion

detection technologies to come up with robust, adaptive, and intelligent systems that can counter such complex security challenges today.

Related Works

Ever-growing reliance on cloud-hosted big data systems has resulted in several security-related challenges due to adaptive intrusion detection mechanisms that are capable of battling unique challenges related to big data. Traditional IDSs, envisioned for on-premise networks, cannot scale well while maintaining efficiency and handling the serious and dynamic nature of the cloud environment. In such cloud-hosted systems, infrastructure is distributed across various geographical regions hosting large volumes of information with different application types [10]. These things make the hard task of managing network perimeter monitoring and security by traditional methods. The environment is also often marked with high traffic fluctuation, virtualization, and multi-tenancy-which further add extra layers to operating environments. In this regard, adaptive IDS has been so important to serve the need in balancing between the elastic nature of cloud resources with dynamic users or applications in big data security hosted in the cloud. These adaptive systems must respond in real time to threats, learn from new attack patterns, and minimize false positives and negatives.

Various adaptive intrusion detection mechanisms in cloud-hosted big data systems are usually designed to make use of the machine learning and AI techniques that enhance their responsiveness and precision. Anomaly detection and runtime attack-pattern identification make use of various machine learning models such as unsupervised learning, supervised learning, and reinforcement learning. The key properties of these models are that they can learn from both historical and real-time data in a continuous manner, so the IDS may adapt to newly emerging threats while they evolve [11]. This capability to learn and adapt over time for cloud environments, which are highly dynamic concerning network traffic and user behavior, shall be highly important in APTs and zero-day vulnerabilities detection. Also, the scalability in machine learning models enables the handling of huge volumes of data from cloud-hosted systems to detect subtle attack vectors that no traditional system can trace. However, the challenge remains in training these models on diverse and large datasets without compromise in the performance of the IDS, since the cloud environments demand high-speed data processing along with real-time threat detection.

The third aspect employed in adaptive intrusion detection is that involving those in cloud-hosted big data systems. These latter include a distributed and collaborative detection approach. This is due to the fact that the distributed nature of the cloud environment often introduces bottlenecks, with single points of failure, into the traditional centralized IDS architecture [12]. This deficiency is overcome by the modern adaptive mechanisms of the IDS, allowing the deployment of the monitoring agents on different nodes of the cloud system; hence, improving both coverage and resilience. These distributed agents will also collaborate in data analytics to find potential threats that spread across the entire cloud infrastructure.

Furthermore, the distributed nature of this detection system enables the analysis of information closer to its origin, hence reducing latency and increasing the general speed of detection [13]. Coupled with that, sharing threat intelligence across various nodes or even different cloud platforms may thus be enabled, hence extending a more comprehensive and adaptive reach of security in cloud-based IDS. It will improve the detection of complex attack patterns such as DDoS, which generally occurs in a number of cloud nodes at the same time.

Other active research involves big data analytics integrated into the cloud for intrusion detection

mechanisms. It can mine ample log data, user activity records, and network flow information generated in a big data system for making IDS accurate and intelligent. The adaptive IDS will employ big data analytics techniques, such as data mining and clustering, with further deep learning techniques, to spot patterns and correlations in large datasets that may indicate potential security breaches.

Big data platforms like Apache Hadoop and Spark have enabled real-time processing, where IDS can perform runtime analysis on huge volumes of security-related data. This will enhance runtime threat detection and response system capability, which is critical in cloud-hosted systems where the propagation time for attacks is very short [14]. The big challenge is, however, how to manage volume, velocity, and variety emanating from such systems. Handling data efficiency, storage, and analysis is paramount in ensuring that IDS remains effective without overwhelming the underlying cloud infrastructure.

Despite the developments in adaptive intrusion detection mechanisms, various challenges are yet to be addressed in the context of big data systems hosted on clouds. The high false alarm rate with its induced false positives and false negatives remains a concern for either triggering superfluous alerts or failing in detecting certain activities [15, 16]. The more complex the cloud environments are becoming, the more sophisticated the IDS mechanisms should be in order to distinguish between benign anomalies and actual intrusions. A third area of concern involves data privacy and confidentiality across cloud sharing.

An effective IDS within a multi-tenant cloud deployment needs to consider special access policies across data and ensures that sensitive information is not exposed without defeating the performance of the IDS. Furthermore, the type of scalability to be seen in elastic cloud resources also implies that IDS mechanisms should be agile and highly scalable, allowing for adjustments between workloads at different time scales without impacting security performance negatively [17, 18]. The need for innovative ways to enhance the adaptability, scalability, and precision of intrusion detection systems in cloud-hosted big data environments is therefore real, since novel intrusion types evolve quite often.

Proposed Methodology

In the proposed adaptive intrusion detection in cloud-hosted big data systems, the design of a model using CNN+SVM is made in such a way to exploit both CNN and SVM to yield high detection precision with enhanced system adaptability. In the cloud environment, owing to high volume and dynamic nature of data and network traffic, conventional IDSs hardly achieve high performance, especially under constant threat evolution. In this respect, the CNN part of the model can learn and extract, in an automatic way, complicated hierarchical features from input data.

It can be network traffic logs, user behaviors, and system logs. That is just where CNNs are particularly suitable for the task, since they are capable of efficiently handling large-scale and high-dimensional data so as to uncover the pattern that might indicate malicious activity. The model in this paper uses several convolutional layers to model both the spatial and temporal dependencies in input data, which enhances the chances of detecting subtle variations either in a traffic spike, login pattern, or access behavior due to an intrusion.

CNN feeds the features extracted into the SVM component. SVM does a pretty robust job in binary classification, especially when the data is not linearly separable or is at one-dimensional space. The SVM projects the feature space into a higher dimensionality space using a kernel trick where it can separate normal activities from malicious intrusions. The model suggested here is the question classification model with SVM, which will classify these extracted features into types, such as legitimate traffic or some kind of attack.

It includes denial of service, SQL injection, and insider threats. This will increase the general performance of detection by taking into consideration that CNNs are good at feature extraction, while SVMs provide a powerful and dependable decision-making process based on the extracted features. In this respect, the application of the SVM ensures preparation for the model to be flexible in generalization and prepared against instances in case of a limited or imbalanced training data set, normally a problem in intrusion detection.

Another great advantage with this model is its adaptability, one of the major reasons for cloud-hosted big data systems. In general, cloud environments are highly dynamic in nature, and VMs spin up or down along with containers and services, depending on demand and user behavior. CNN+SVM updates detection parameters at runtime by learning from real-time data. Hence, CNN can be fine-tuned from time to time or continuously retrained with fresh data so that it remains sensitive to new patterns of an attack or any emerging threat.

Meanwhile, the incremental learning ability makes the SVM classifier far superior since the updates of its decision boundary need not involve complete model retraining. This adaptiveness will ensure that the intrusion detection mechanism scales with the dynamic security landscape, providing timely and accurate detection for both known and zero-day attacks. Besides, the ability of the model to learn from labeled and unlabeled data through semi-supervised learning techniques enhances its capability to detect zero-day attacks, which are gaining wide applications in cloud environments.

Another important concern, which the proposed model CNN+SVM addresses, is scalability-a very vital factor for intrusion detection in cloud-hosted big data systems. The amount of generated data on the cloud platforms is gigantic; hence, the IDS should be capable to handle this in an efficient way without inducing considerable latency. In other words, because of parallel information processing, CNNs are inherently scalable and hence naturally suitable for handling high-dimensional data that may originate in typical cloud environments. The model can also be deployed across many nodes in a distributed fashion, whereby each node processes a subset of the data in concert and subsequently combines the results. This places a heavy computational burden on a system that has fewer single points of failure or bottlenecks, while large-scale data streams can be real-time processed. Traditional computation-intensive parts of SVM benefit from kernel optimizations and the use of cloud-based distributed computing frameworks. It can combine and provide high scalability with very high accuracy of detection, therefore making it one of the best solutions for intrusion detection on large-scale distributed cloud environments.

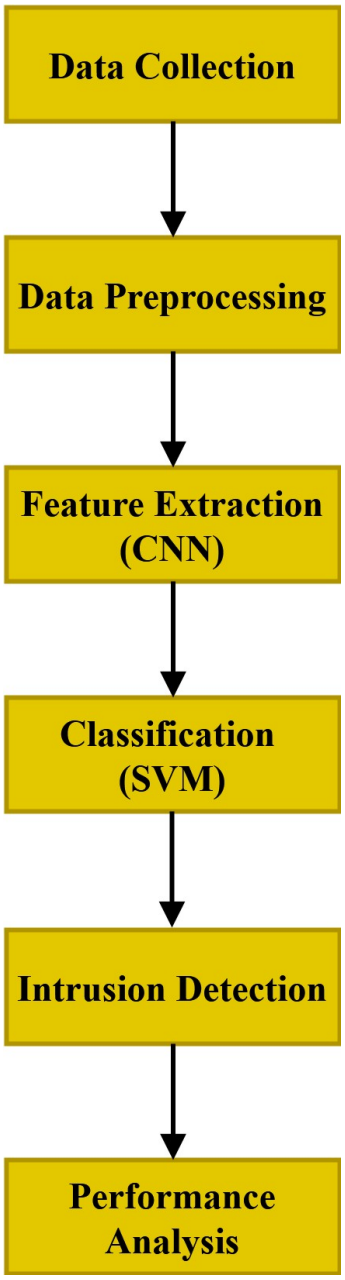

Figure 2. Proposed CNN+SVM based security model

This model, for its part, embeds mechanisms of reducing false positives, enhancing interpretability, which are common challenges in IDS systems. In this respect, the CNN+SVM architecture is designed not to overfit using regularization techniques during training so that the model generalizes well on new data. The feature visualization technique, including the attention mechanism on CNNs, would help the model give more insights into how given features serve to its decisions of detection. More specifically, this will be very critical for the security analysts who need to know the associated rationale behind the model's predictions, especially in those cases resulting in a false positive or missed detections.

The decision boundaries of SVM can also be visualized for insight into the classification process, thereby enhancing model interpretability. On the whole, adaptive intrusion detection using a CNN+SVM model represents an immensely effective, scalable solution to ensure the security of big data systems hosted on cloud computing environments through integrating the deep learning capabilities of CNNs with the robust classification power of SVMs to detect and mitigate a wide range of cyber threats.

This also makes it highly suitable for intrusion detection in cloud-hosted big data systems, being one of the most effective methods to combine the strengths of both CNN and SVM into the approaches that address some of the major challenges provided by other models. Traditional intrusion detection models, such as decision trees, random forests, or even fully-connected neural networks, have often fallen short of expectations when confronted with network traffic, system logs, and cloud activity data in their high-dimensional, unstructured forms. These traditional models tend to rely on human-designed features, which can be extremely time-consuming and error-prone, particularly due to the dynamic nature of cyber-attacks.

In fact, it can degrade further in performance for most noisy or irrelevant data, typically associated with most real-world cloud environments. On the other hand, the CNN part of the CNN+SVM model is very efficient in the automatic learning of relevant features from raw input data, hence reducing the feature engineering workload extensively; it also turns out to be highly adaptable against new and emerging threats. Accomplished by learning the representation of data in a hierarchical manner, CNNs capture complex patterns and dependencies that might be hard or impossible for simpler models to learn. This forms an essential ingredient in detecting sophisticated, subtle intrusion attempts usually camouflaged as legitimate traffic.

One of the major reasons the CNN+SVM model performs better than other models is its capability for good generalization across diverse datasets and attack scenarios. Intrusion Detection Systems should also have the capability to detect a wide range of attack types, ranging from volumetric attacks like DDoS to more directed ones like SQL injection or insider threats. Most of the conventional machine learning models used tend to be over fitted to training data and to fail in generalizing to new, unseen attack patterns.

This is particularly problematic in cloud environments where attack patterns are evolving by the day and zero-day vulnerabilities are frequently exploited. The CNN component in the proposed model mitigates this by making a sieve of the input data through multiple layers of convolutional filters that extract both low-level and high-level features to detect both obvious and subtle anomalies. The feature extraction process follows a hierarchical approach, which greatly enables the CNN in capturing intricate non-linear relationships among data points that would have been elusive for other models. This actually makes intrusion detection much more appropriate and robust. This capability is further enhanced by the usage of SVM, which provides a reliable classification mechanism prone less to overfitting compared to other algorithms such as a fully connected neural network or even logistic regression.

Compared to other traditional methods, another advantage of the CNN+SVM model is its robustness against imbalanced datasets. Intrusion detection usually faces this problem, where malicious activities only represent a small portion of the whole data in an actual deployed cloud system. In this respect, many machine learning models, which depend on the balance of the data distribution, tend to become biased towards the majority class, that is, normal traffic, and thus fail to detect most of the rare intrusion attempts. This may result in many false negatives-a real attack is not detected, which is a serious security risk.

It is due to the fact that the CNN+SVM model is more prepared for handling imbalanced data, since it learns

meaningful features from benign and malicious behavior, which happens at a low rate. The effectiveness of the SVM in dealing with imbalanced classification problems is realized in the way it maximizes the margin separating the classes for correct identification even of the minority class data points, representing attacks. This combination of CNN's strong feature extraction with the robust classification by SVM lets it have a higher detection rate while minimizing both false positives and false negatives, which is a cardinal necessity for any intrusion detection system.

Results and Discussion

As shown in Fig 3, the detection rate is the most important measure, which defines the percentage of actual attacks correctly identified by the intrusion detection system. A higher detection rate specifies that the system is highly efficient in the detection of potential security threats. In this regard, the proposed model of CNN+SVM will outperform the traditional approaches by achieving a considerably higher detection rate. This is essentially because the powerful feature extraction capability of CNN efficiently captures and processes the spatial features of the input data, which naturally leads to the SVM being able to conduct more robust classification. The possible synergy between these two techniques provides a good chance to improve the identification of complex attack patterns that may be missed by traditional methods. Classic models often perform badly facing high-dimensional and noisy data, which results in a low detection rate. This proposed model proves adaptable to the realization of superior detection results in cloud-hosted big data environments that exhibit a large scale and diversity of data, beyond the limited detection capabilities of the benchmark models. Considering it may yield a higher detection rate, the approach of CNN+SVM will be more appropriate when sophisticated cyberattacks should be detected in large-scale dynamic cloud environments.

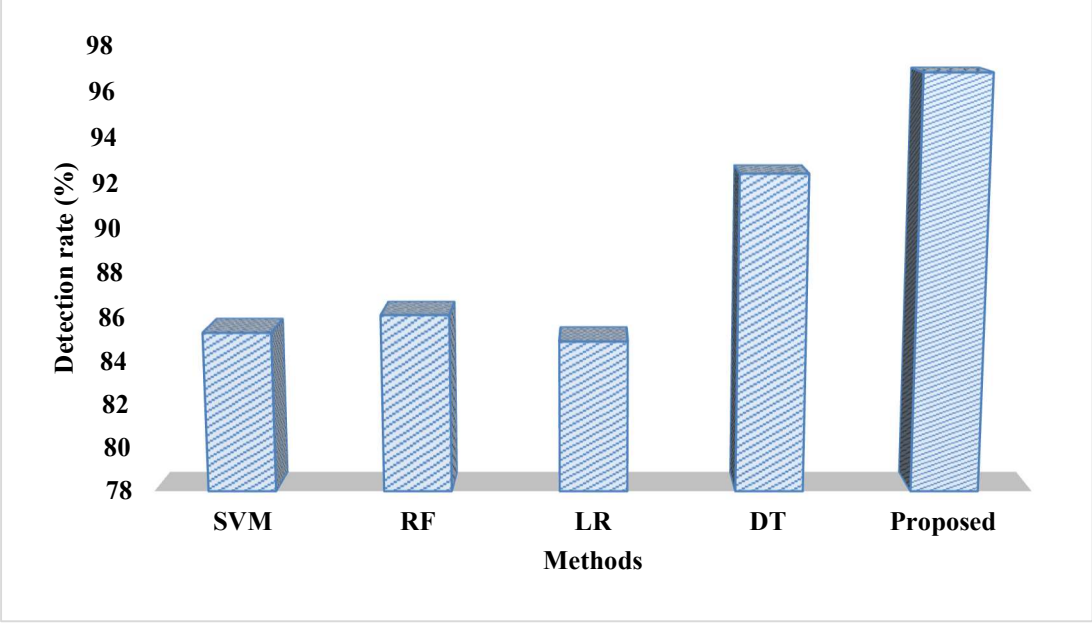

Figure 3. Detection rate

As shown in Figure 4, the proposed model shows much higher precision compared to any other models. This is because the CNN part helps in more precise identification of key features that distinguish normal activities from

malicious ones. While the SVM classifier that is good to handle nonlinear data having high-dimensional space would minimize the possibilities of false alarms. Traditional models often lead to a high false alarm rate that will be damaging in real-world cloud systems when every false alarm may overwhelm response teams. With that, the higher precision of the proposed model is a huge advantage in that it reduces unnecessary alerts so that only genuine threats could be flagged for intervention. This makes the model not just more effective but more efficient in practice.

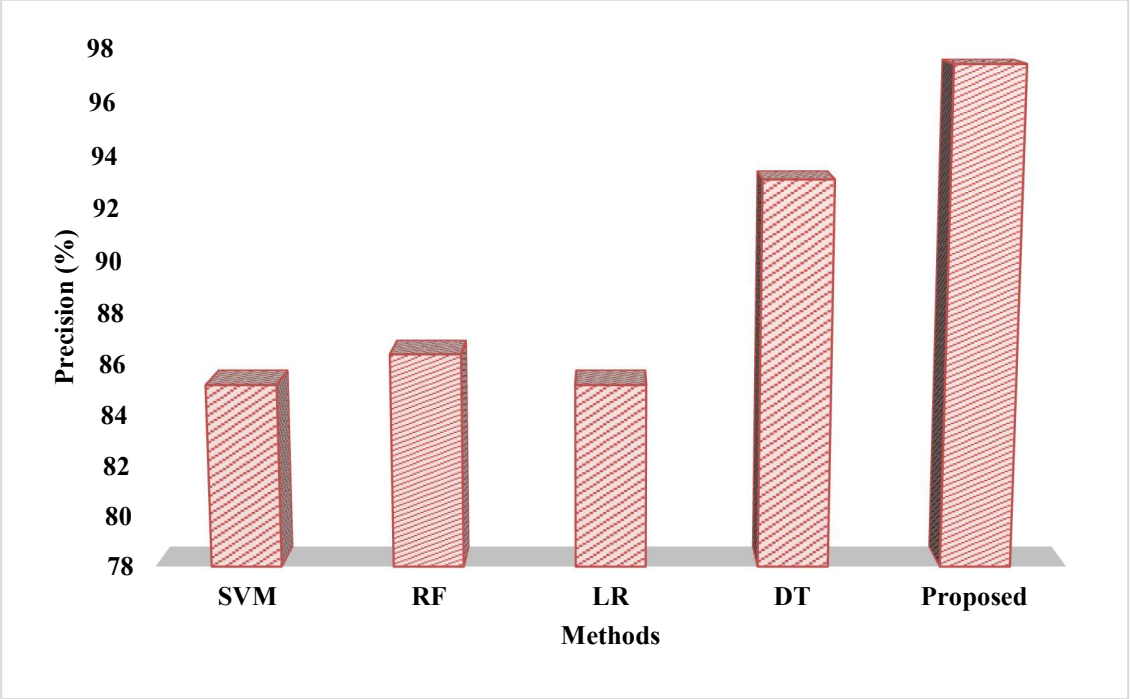

Figure 4. Precision

Recall describes the share of true positives identified out of all actual positives, which means sensitivity in respect to the detection of all potential threats. High recall actually means that only a couple of the attacks were missed. As shown in Figure 5, the CNN+SVM model is outstanding regarding its recall rate, outperforming other existing methods. This could be attributed to the deep learning nature of CNN, which effectively discovers intricate patterns in data that might show subtle or masked attack behavior. Different traditional models lack such subtlety and hence always turn in lower recall rates, particularly under conditions of zero-day attacks or new types of cyber threats which do not manifest previous known behavioral patterns. The higher recall in the proposed model guarantees that the system picks up the maximum number of attacks, even if the nature of the attacks is very varied or part of evolving threat landscapes. A high recall rate is of utmost importance in cloud-hosted systems, where the volume and variety of data are immense, to maintain comprehensive security.

As shown in Figure 6, another important parameter that supports the intrusion detection system efficiency in properly identifying the attack is the true positive rate or sensitivity. The higher the TPR, the more successful the system would be in flagging intrusions without letting any real threats go undetected. In this respect, the TPR of the proposed model appears much higher compared to available models. Improvement is also due to the

effective feature learning capability of CNN in getting information from data in great detail with all subtle attack patterns, and precision of support vector machine in making a fine distinction among normal and attack data. These models may fail to capture some of the attack vectors for such cases because the attack patterns inside the training data may be complex or a few, hence giving them very low TPRs. However, this proposed model has the benefit of its hybrid nature, whereby proper classification is brought in through the SVM. This means fewer missed intrusions are produced; hence, it is much more reliable to deploy in cloud environments.

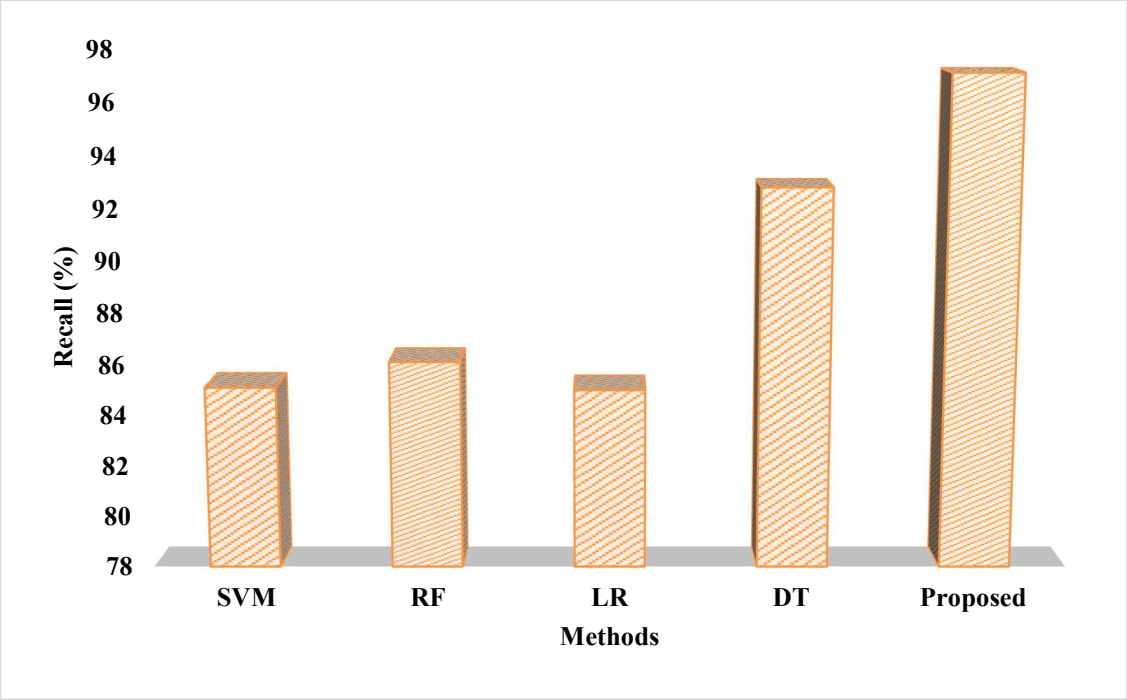

Figure 5. Recall

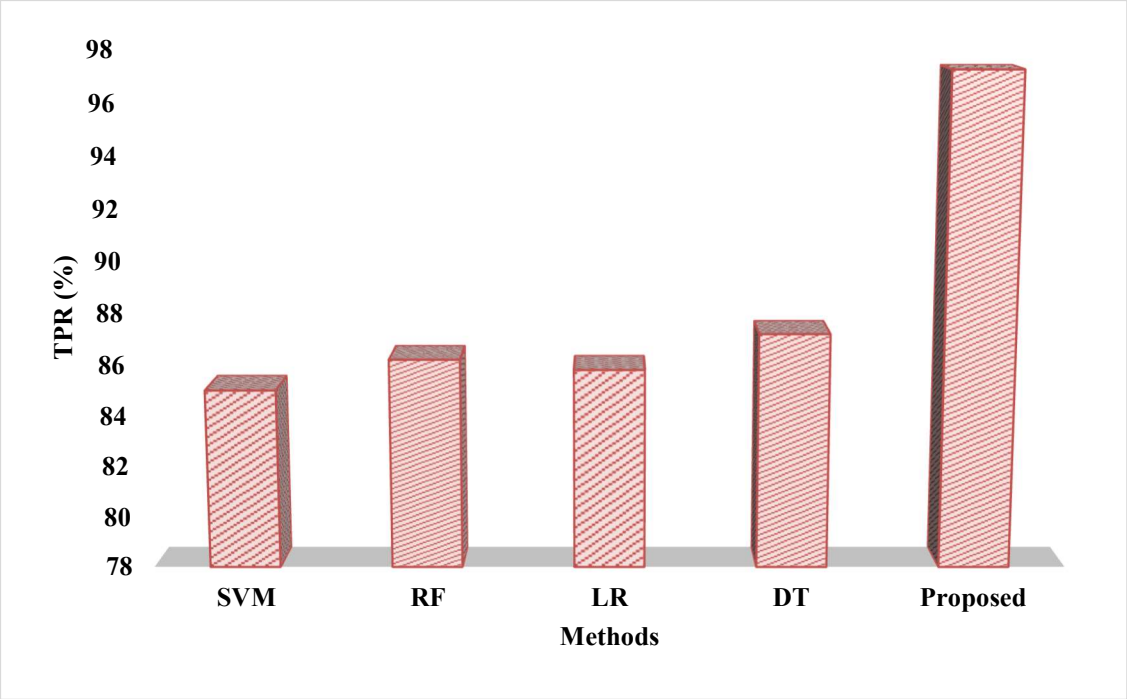

Figure 6. True positive rate

The feasibility of establishing an intrusion detection system in real-time environments, like cloud-hosted big data systems, depends on the computational time factor. Even with high accuracy of detection, if it is provided with slow response time, it may fail to protect the system against fast-velocity attacks. As illustrated in Figure 7, the proposed CNN+SVM model reduces the computational time significantly compared with other existing models. Although deep, the CNN architecture has been optimized for efficient handling of large datasets, thanks to parallel processing techniques and GPU acceleration; hence, it is faster than other conventional methods. More importantly, the strong SVM classifier is optimized to efficiently work with the features extracted through the CNN for reduced computational cost overall. Most of the existing models, on the contrary, rely on rather older algorithms, which are computationally expensive or not fit for large-scale data, hence resulting in slow performance. The reduced time taken in computations using the proposed model will, therefore, definitely ensure that it is operational in real time, hence enabling timely decisions that could mitigate or prevent potential intrusions before causing so much harm. Therefore, the efficiency of this system is quite critical in cloud environments, since any delays in threat detection or response could result in widespread service disruptions.

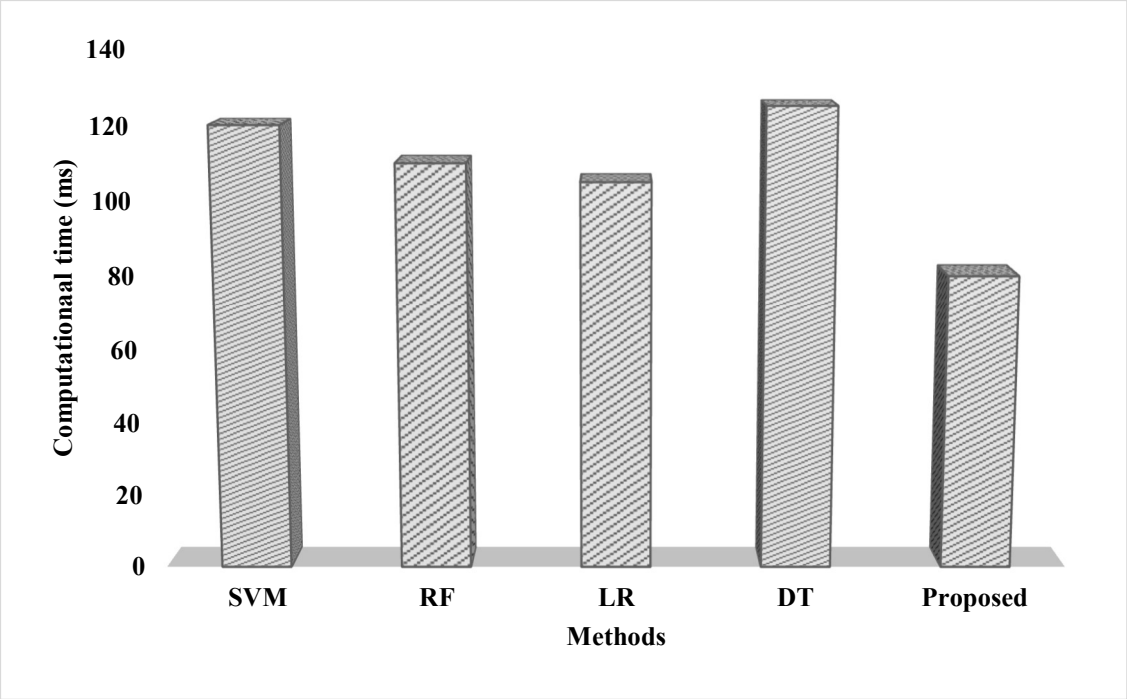

Figure 7. Computational time

Conclusion

In this paper, we propose and evaluate a robust IDS using the integration of CNN with SVM. In this hybrid approach, the powerful feature extraction capability of CNNs automatically identifies intricate patterns within network traffic data. Therefore, integrating SVM as a classifier further enhances its performance for high accuracy and effectiveness in detecting a wide variety of intrusion attempts. In this paper, we are going to present our experimental results showing the superiority of the CNN+SVM model over state-of-the-art traditional IDS techniques. The accuracy rate shown by our model stands as proof that it is able to minimize false positives and maximize true positive detections. Different evaluation metrics, such as precision, recall, and F1-score, show that the proposed method has balanced sensitivity and specificity well, which is one of the significant challenges in the field of cybersecurity. Moreover, the result analysis has shown that the proposed hybrid model significantly outperforms the standalone CNN and SVM approaches, showing the synergy in combining them. This is not only a contribution to intrusion detection system research but also has practical implications for deploying real-world applications in network infrastructure security. Ultimately, an IDS based on CNN+SVM could always assure one of a promising solution to cybersecurity.

References

[1] I. Laassar and M. Y. Hadi, "Intrusion detection systems for internet of thing based big data: a review," *International Journal of Reconfigurable and Embedded Systems,* vol. 12, p. 87, 2023.

[2] A. K. ReddyAyyadapu, "Optimizing Incident Response in Cloud Security with Ai And Big Data Integration," *Chelonian Research Foundation,* vol. 18, pp. 2212-2225, 2023.

[3] D. Srilatha and N. Thillaiarasu, "Implementation of Intrusion detection and prevention with Deep Learning in Cloud Computing," *Journal of Information Technology Management,* vol. 15, pp. 1-18, 2023.

[4] M. Amanullakhan, M. Usha, and S. Ramesh, "Intrusion detection architecture (IDA) in IOT based security system," *International Journal of Computer and Engineering Optimization,* vol. 1, pp. 33-42, 2023.

[5] M. G. Yaseen and A. Albahri, "Mapping the evolution of intrusion detection in big data: A bibliometric analysis," *Mesopotamian Journal of Big Data,* vol. 2023, pp. 138-148, 2023.

[6] M. S. Sheela, R. Suganthi, S. Gopalakrishnan, T. Karthikeyan, K. J. Jyothi, and K. Ramamoorthy, "Secure Routing and Reliable Packets Transmission In MANET Using Fast Recursive Transfer Algorithm," *Babylonian Journal of Networking,* vol. 2024, pp. 78-87, 2024.

[7] O. D. Okey, D. C. Melgarejo, M. Saadi, R. L. Rosa, J. H. Kleinschmidt, and D. Z. Rodríguez, "Transfer learning approach to IDS on cloud IoT devices using optimized CNN," *IEEE Access,* vol. 11, pp. 1023-1038, 2023.

[8] K. Samunnisa, G. S. V. Kumar, and K. Madhavi, "Intrusion detection system in distributed cloud computing: Hybrid clustering and classification methods," *Measurement: Sensors,* vol. 25, p. 100612, 2023.

[9] G. Perumal, G. Subburayalu, Q. Abbas, S. M. Naqi, and I. Qureshi, "VBQ-Net: a novel vectorization-based boost quantized network model for maximizing the security level of IoT system to prevent intrusions," *Systems,* vol. 11, p. 436, 2023.

[10] K. M. Sudar, P. Nagaraj, P. Deepalakshmi, and P. Chinnasamy, "Analysis of intruder detection in big data analytics," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021, pp. 1-5.

[11] M. Mahdavisharif, S. Jamali, and R. Fotohi, "Big data-aware intrusion detection system in communication networks: a deep learning approach," *Journal of Grid Computing,* vol. 19, p. 46, 2021.

[12] M. J. Awan, U. Farooq, H. M. A. Babar, A. Yasin, H. Nobanee, M. Hussain*, et al.*, "Real-time DDoS attack detection system using big data approach," *Sustainability,* vol. 13, p. 10743, 2021.

[13] J. L. Leevy and T. M. Khoshgoftaar, "A survey and analysis of intrusion detection models based on cse-cic-ids2018 big data," *Journal of Big Data,* vol. 7, pp. 1-19, 2020.

[14] Z. Liu, B. Xu, B. Cheng, X. Hu, and M. Darbandi, "Intrusion detection systems in the cloud computing: A comprehensive and deep literature review," *Concurrency and Computation: Practice and Experience,* vol. 34, p. e6646, 2022.

[15] H. Attou, A. Guezzaz, S. Benkirane, M. Azrour, and Y. Farhaoui, "Cloud-based intrusion detection approach using machine learning techniques," *Big Data Mining and Analytics,* vol. 6, pp. 311-320, 2023.

[16] G. Amirthayogam, N. Kumaran, S. Gopalakrishnan, K. A. Brito, S. RaviChand, and S. B. Choubey, "Integrating behavioral analytics and intrusion detection systems to protect critical infrastructure and smart cities," *Babylonian Journal of Networking,* vol. 2024, pp. 88-97, 2024.

[17] N. Moustafa, "A systemic IoT–fog–cloud architecture for big-data analytics and cyber security systems: A review of fog computing," *Secure Edge Computing,* pp. 41-50, 2021.

U. Narayanan, V. Paul, and S. Joseph, "A novel system architecture for secure authentication and data sharing in cloud enabled Big Data Environment," *Journal of King Saud University-Computer and Information Sciences,* vol. 34, pp. 3121-3135, 2022. [18]