Research Article

ACCENTS Transactions on Information Security, Vol 9(35)
ISSN (Online): 2455-7196
http://dx.doi.org/10.19101/TIS.2024.935002

Evolution and advancements in intrusion detection systems: from traditional methods to deep learning and federated learning approaches

Ashish Kumar Ranjan* **and Animesh Kumar Dubey**
Department of Computer Science and Engineering, Patel Collage of Science and Technology, Bhopal (M.P) India

Received: 10-March-2024; Revised: 03-July-2024; Accepted: 06-July-2024

Abstract
Intrusion Detection Systems (IDS) are crucial for maintaining the security and integrity of network infrastructures. This review paper comprehensively examines the evolution and advancements in IDS technologies, focusing on both traditional methods and contemporary machine learning and deep learning approaches. Traditional IDS methods, including signature-based and anomaly-based detection, laid the groundwork for current systems but faced challenges such as high false-positive rates and limited adaptability. Recent advancements in machine learning, specifically supervised and unsupervised learning algorithms, have significantly enhanced the accuracy and efficiency of IDS. Deep learning techniques, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), offer promising solutions for complex and high-volume network traffic analysis. This review also explores federated learning for IDS, emphasizing its potential for data privacy preservation and reduced computational load. Furthermore, hybrid models combining various algorithms are discussed for their capability to leverage the strengths of multiple techniques. The paper synthesizes current research, highlighting key methodologies, datasets, evaluation metrics, and the future direction of IDS research. By providing a thorough analysis of existing literature and identifying gaps, this review aims to guide future research efforts and practical implementations in the field of network security.

Keywords
Intrusion detection system (IDS), Machine learning, Deep learning, Network security, Federated learning.

1.Introduction
In the modern digital age, the security of network infrastructures is of paramount importance. The rapid proliferation of internet-connected devices and the increasing complexity of network environments have amplified the necessity for robust security mechanisms [1, 2]. Intrusion detection systems (IDS) play a vital role in safeguarding networks against malicious activities, unauthorized access, and cyberattacks [3–5]. IDS monitor network traffic, analyze data for suspicious patterns, and alert administrators to potential security breaches, thereby acting as a critical line of defense in cybersecurity strategies [6, 7]. The evolution of IDS technology has been marked by significant advancements, beginning with traditional methods and progressing to sophisticated machine learning and deep learning approaches [8–10]. Early IDS implementations were primarily signature-based, relying on predefined patterns of known threats to detect intrusions [7–10].

While effective against known attacks, these systems were limited by their inability to identify novel threats and their susceptibility to high false-positive rates [6-10]. Anomaly-based detection methods emerged as an improvement, leveraging statistical models and behavioral analysis to identify deviations from normal network activity. However, these approaches also faced challenges in accurately distinguishing between benign anomalies and actual threats [7–12].

The advent of machine learning introduced a paradigm shift in IDS development. ML algorithms, with their ability to learn from data and improve over time, offered promising solutions for enhancing detection accuracy and reducing false positives. [11, 12] Supervised learning techniques, such as decision trees (DT), support vector machine (SVM), and k-nearest neighbors (kNN), were applied to classify network traffic based on labeled training data. Unsupervised learning methods, including clustering and anomaly detection algorithms, enabled the identification of previously unknown attack patterns without the need for labeled data. Despite their advancements, traditional ML approaches still encountered

*Author for correspondence

difficulties in handling the high dimensionality and dynamic nature of network traffic.

Deep learning DL, a subset of ML, has revolutionized the field of IDS with its powerful capabilities in feature extraction and pattern recognition. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been particularly successful in analyzing complex network traffic data [4–8]. CNNs excel in identifying spatial features within traffic flows, while RNNs capture temporal dependencies, making them suitable for detecting sequential attack patterns. Hybrid models combining CNNs and RNNs have demonstrated superior performance in various intrusion detection tasks.

Federated learning represents a recent innovation in IDS research, addressing concerns related to data privacy and computational efficiency [10–12]. By enabling collaborative model training across multiple decentralized devices without sharing raw data, it preserves user privacy while leveraging the collective intelligence of distributed networks. *Figure 1* highlights the methods considered for review and analysis.

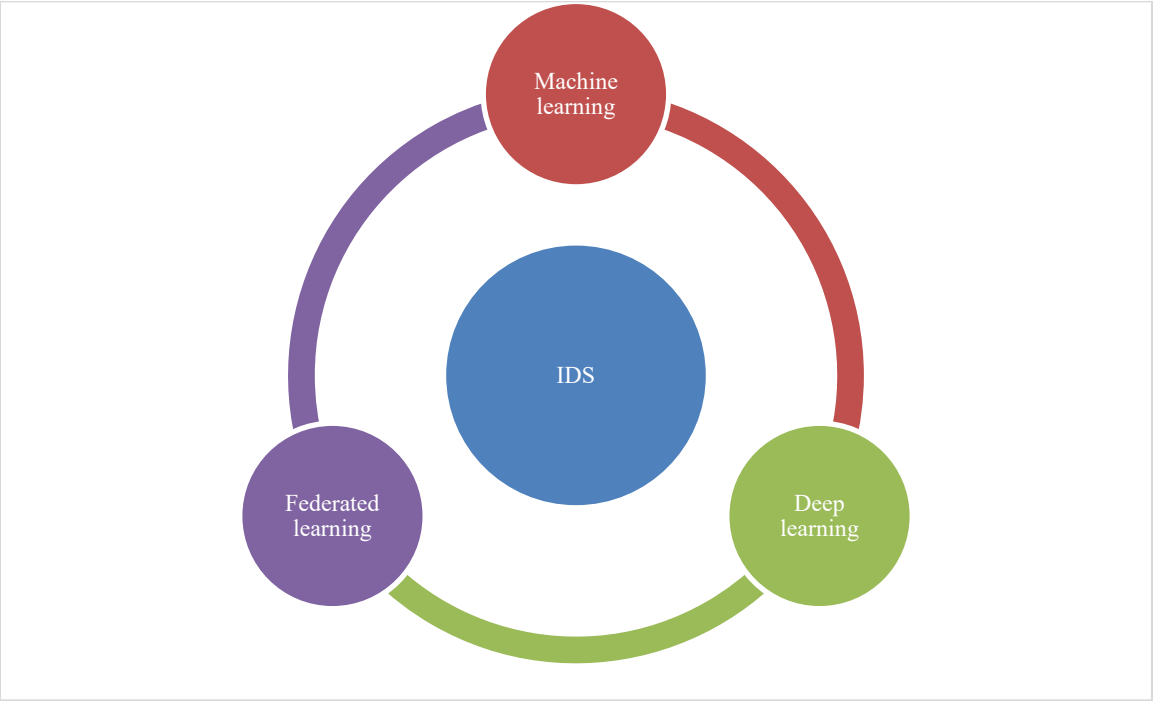

Figure 1 Approaches covered for the review and analysis

This review paper aims to provide a comprehensive analysis of the current state of IDS technology, highlighting the evolution from traditional methods to advanced machine learning and deep learning approaches. Key methodologies, evaluation metrics, and notable datasets used in IDS research will be discussed. Furthermore, the paper will explore emerging trends and future directions in IDS development, identifying challenges and opportunities for further advancements. By synthesizing existing literature and identifying research gaps, this review seeks to inform and guide future research efforts, ultimately contributing to the development of more effective and resilient intrusion detection systems.

This paper is organized as follows. Related work has been explored in section 2. Section 3 covers the discussion and analysis and finally it is concluded in section 4.

2.Literature review

In 2023, Nivedhidha et al. [13] highlighted the vulnerability of highly interconnected hospital networks to cyber threats, emphasizing the necessity of a robust Network Intrusion Detection System (NIDS). To combat the challenges posed by class imbalance in network intrusion data, they proposed a machine learning-based NIDS utilizing CopulaGAN to generate synthetic samples and balance the dataset. This approach, combined with a finely tuned random

forest (RF) algorithm, significantly enhances the detection accuracy, making it an effective solution for safeguarding hospital networks.

An ImprovedFedAvg algorithm was proposed for network intrusion detection, enhancing federated learning for better data privacy and reducing transmission of model weights (Lee et al. (2023)) [14]. This addressed the security and legal concerns associated with centralized data collection. The approach leveraged deep learning to improve model performance in detecting network intrusions.

A CNN model was developed for network intrusion detection, addressing dimension explosion with an albino principal component analysis (PCA) algorithm (Zhang et al. (2023)) [15]. Dropout layers were introduced to prevent overfitting, using Sigmoid for classification. The proposed method demonstrated improved accuracy and robustness in intrusion detection.

Feng et al. (2023) [16] introduced IP-filtered multi-channel convolutional neural network (IP-MCCLSTM) to enhance network intrusion detection efficiency. By filtering traffic by IP, system loading was reduced. The method achieved 98.9% accuracy and 99.7% Macro-Recall on the 2017CICIDS dataset, outperforming other comparison methods.

Roy et al. (2023) [17] utilized a feature selection method and fully convolutional network architecture. It was developed for software-defined networking (SDN), addressing vulnerabilities in wireless sensor networks. The approach minimized data volume relayed and improved detection accuracy, as validated using the UNR-IDD dataset.

A hybrid intrusion detection model combining CNN, bidirectional long short-term memory (BiLSTM), and attention mechanisms was proposed for SDN (Said and Askerzade (2023)) [18]. The model effectively captured complex network traffic patterns and demonstrated high accuracy in detecting various intrusions. Evaluation on the InSDN dataset showed superior performance compared to state-of-the-art models.

A neural network-based attack classifier was developed using the advanced security network metrics dataset (Desai and Gopalakrishnan (2023)) [19]. Feature selection techniques, such as Variance threshold and Chi-square Test, were applied to enhance classification accuracy. The resulting model achieved a 99% prediction accuracy in network intrusion detection.

Lonare et al. (2024) [20] conducted an analysis of NIDS methodologies, algorithms, and technologies was conducted. The paper categorized NIDS into signature-based, anomaly-based, and hybrid approaches, evaluating their strengths and weaknesses. Various techniques were discussed, and scalable k-means with RF, to provide insights for cybersecurity advancements.

An intrusion detection method based on payload analysis using a Transformer and BiLSTM network was proposed (Wanshun et al. (2023)) [21]. This approach focused on extracting byte-level and packet-level features for better detection of malicious traffic. Experiments with the CICIDS2017 dataset showed improved performance compared to other methods.

A deep learning-based computer network intrusion detection method was proposed (Huang (2023)) [22], addressing limitations of traditional methods in accuracy and speed. Experiments demonstrated that the new approach achieved better detection accuracy and lower false positive rates, making it suitable for real-time network intrusion detection.

A variational auto-encoder and generative adversarial network were used to address imbalanced attack traffic classification in network intrusion detection (Lu and Jiao (2023)) [23]. The method generated diverse sample classes to balance training data and improve unknown attack detection. Validation showed the scheme's effectiveness in enhancing detection capabilities.

3.Discussion and analysis

IDS have undergone significant transformations over the past decades, driven by the need to safeguard increasingly complex network infrastructures. Traditional IDS methods, while foundational, have shown limitations in scalability and adaptability. Signature-based detection, for instance, relies heavily on predefined attack patterns, making it ineffective against novel threats. Anomaly-based detection, although more adaptive, often struggles with high false-positive rates due to its reliance on deviations from established norms.

Integrating machine learning into IDS has improved accuracy in detecting known attack patterns using supervised learning like DT and SVM, though they require extensive labeled data. Unsupervised methods

identify new attacks but may produce false alarms. Deep learning, using CNNs and RNNs, enhances detection but is complex and resource-intensive. Federated learning addresses privacy and computational concerns by enabling decentralized model training.

Despite these advancements, several challenges persist. Achieving real-time processing with minimal latency remains a critical hurdle, especially as network traffic volume continues to grow. Moreover, the resilience of IDS to adversarial attacks, where attackers manipulate input data to evade detection, is an ongoing area of concern. Ensuring the robustness and reliability of IDS in diverse and evolving threat landscapes requires continuous innovation and refinement.

While significant progress has been made in IDS technology, ongoing research and development are essential to address current limitations and enhance the efficacy of these systems. The integration of advanced machine learning and deep learning techniques, coupled with innovations in FL, holds promise for the future of IDS. However, achieving a balance between detection accuracy, computational efficiency, and privacy preservation remains a complex and critical challenge.

4.Conclusion

IDS have become an indispensable component of network security, evolving significantly from traditional signature-based and anomaly-based methods to sophisticated machine learning and deep learning approaches. This review has highlighted the strengths and limitations of various IDS methodologies, including supervised and unsupervised learning techniques, and the revolutionary impact of deep learning models such as CNNs. Additionally, the integration of federated learning has shown potential in addressing privacy concerns and enhancing model performance through decentralized data processing.

Despite these advancements, challenges remain in developing IDS that can effectively handle the increasing complexity and volume of network traffic. The dynamic nature of cyber threats necessitates continuous updates and improvements in IDS algorithms to maintain their efficacy. Moreover, achieving a balance between detection accuracy and computational efficiency is critical to deploying IDS in real-world scenarios.

Some of the future works are as under:

1. Future work should focus on developing advanced feature extraction and selection techniques to improve IDS's ability to distinguish between normal and malicious traffic.
2. Future research should explore hybrid approaches that combine multiple machine learning and deep learning techniques to leverage their complementary strengths. Developing ensemble models that integrate outputs from various algorithms can enhance detection accuracy and reduce false positives.
3. Advancing real-time intrusion detection capabilities should be a priority, focusing on optimizing algorithms for faster processing and minimizing latency without compromising detection performance.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Muneer S, Farooq U, Athar A, Ahsan Raza M, Ghazal TM, Sakib S. A critical review of artificial intelligence based approaches in intrusion detection: a comprehensive analysis. Journal of Engineering. 2024;2024(1):3909173.

[2] Amru M, Kannan RJ, Ganesh EN, Muthumarilakshmi S, Padmanaban K, Jeyapriya J, et al. Network intrusion detection system by applying ensemble model for smart home. International Journal of Electrical & Computer Engineering (2088-8708). 2024; 14(3).

[3] Nandanwar H, Katarya R. Deep learning enabled intrusion detection system for industrial IOT environment. Expert Systems with Applications. 2024; 249:123808.

[4] Wang Z, Li J, Yang S, Luo X, Li D, Mahmoodi S. A lightweight iot intrusion detection model based on improved bert-of-theseus. Expert Systems with Applications. 2024; 238:122045.

[5] Turukmane AV, Devendiran R. M-MultiSVM: an efficient feature selection assisted network intrusion detection system using machine learning. Computers & Security. 2024; 137:103587.

[6] Talukder MA, Sharmin S, Uddin MA, Islam MM, Aryal S. MLSTL-WSN: machine learning-based intrusion detection using SMOTETomek in WSNs. International Journal of Information Security. 2024; 23(3):2139-58.

[7] Kumar GS, Kumar RK, Kumar KP, Sai NR, Brahmaiah M. Deep residual convolutional neural network: an efficient technique for intrusion detection system. Expert Systems with Applications. 2024; 238:121912.

[8] Zhang J, Peter JD, Shankar A, Viriyasitavat W. Public cloud networks oriented deep neural networks for

effective intrusion detection in online music education. Computers and Electrical Engineering. 2024; 115:109095.
[9] Karthikeyan M, Manimegalai D, RajaGopal K. Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection. Scientific Reports. 2024; 14(1):231.
[10] Latif S, Boulila W, Koubaa A, Zou Z, Ahmad J. Dtl-ids: an optimized intrusion detection framework using deep transfer learning and genetic algorithm. Journal of Network and Computer Applications. 2024; 221:103784.
[11] Bukhari SM, Zafar MH, Abou Houran M, Moosavi SK, Mansoor M, Muaaz M, et al. Secure and privacy-preserving intrusion detection in wireless sensor networks: Federated learning with SCNN-Bi-LSTM for enhanced reliability. Ad Hoc Networks. 2024; 155:103407.
[12] Chen X, Qiu W, Chen L, Ma Y, Ma J. Fast and practical intrusion detection system based on federated learning for VANET. Computers & Security. 2024; 142:103881.
[13] Nivedhidha M, Ramkumar MP, GSR ES. CopulaGAN boosted random forest based network intrusion detection system for hospital network infrastructure. In 2023 14th international conference on computing communication and networking technologies (ICCCNT) 2023 (pp. 1-7). IEEE.
[14] Lee BS, Kim JW, Choi MJ. Federated learning based network intrusion detection model. In 24st asia-pacific network operations and management symposium (APNOMS) 2023 (pp. 330-3). IEEE.
[15] Zhang P, Tian G, Dong H. Research on network intrusion detection based on Whitening PCA and CNN. In 7th international conference on smart grid and smart cities (ICSGSC) 2023 (pp. 232-7). IEEE.
[16] Feng Q, Lin Z, Bing L. IP-MCCLSTM: a network intrusion detection model based on ip filtering. In 20th international computer conference on wavelet active media technology and information processing (ICCWAMTIP) 2023 (pp. 1-6). IEEE.
[17] Roy B, Acharya I, Papalkar D, Joseph M. Top-performing unifying architecture for network intrusion detection in SDN using fully convolutional network. In 5th international conference on inventive research in computing applications (ICIRCA) 2023 (pp. 1340-4). IEEE.
[18] Said RB, Askerzade I. Attention-based CNN-BiLSTM deep learning approach for network intrusion detection system in software defined networks. In 5th international conference on problems of cybernetics and informatics (PCI) 2023 (pp. 1-5). IEEE.
[19] Desai R, Gopalakrishnan VT. Network intrusion detection through machine learning with efficient feature selection. In 15th international conference on communication systems & networks (COMSNETS) 2023 (pp. 797-801). IEEE.
[20] Lonare MB, Joshi BC, Tripathy SK, Kumar S, Tiwari S. Real-time network monitoring and reporting using network intrusion detection system. In 9th international conference for convergence in technology (I2CT) 2024 (pp. 1-6). IEEE.
[21] Wanshun L, Panxiang Z, Gang D, Min T. BI-TBL: a network intrusion detection method based on payload analysis. In 20th international computer conference on wavelet active media technology and information processing (ICCWAMTIP) 2023 (pp. 1-5). IEEE.
[22] Huang X. Research on computer network intrusion detection algorithm based on deep learning. In IEEE international conference on electrical, automation and computer engineering (ICEACE) 2023 (pp. 1122-5). IEEE.
[23] Lu Y, Jiao P. A classification method for network intrusion detection based on deep generative model. In international conference on mobile internet, cloud computing and information security (MICCIS) 2023 (pp. 162-7). IEEE.

Ashish Kumar Ranjan completed his M.Tech in Computer Science & Engineering at Patel College of Science and Technology, Bhopal (MP), India, and holds a B.Tech degree in Computer Science & Engineering from RTC Institute of Technology, Ranchi, Jharkhand, India. His primary area of interest is Network Security.
Email: ashishcse001@gmail.com

Animesh Kumar Dubey is currently serving as an Assistant Professor in the Department of Computer Science and Engineering at Patel College of Science and Technology, Bhopal, Madhya Pradesh, India. He holds a Bachelor of Engineering (B.E.) and an M.Tech. degree in Computer Science Engineering from Rajiv Gandhi Technical University, Bhopal, Madhya Pradesh. With over 15 publications in reputable, peer-reviewed national and international journals and conferences, his expertise extends across a range of subjects. His primary research interests include Data Mining, Optimization, Machine Learning, Cloud Computing, and Artificial Intelligence.
Email: animeshdubey123@gmail.com