Intelligent Network Intrusion Detection: A Review of Machine Learning and Deep Learning Approaches

Alok kumar[1], Kamlesh Raghuwanshi[2], and Saurabh Karsoliya[3]
[1,2,3] Technocrats Institute of Technology, Bhopal

Abstract. The rapid expansion of digital networks has significantly increased exposure to cyber threats, making effective network security mechanisms indispensable. Network Intrusion Detection Systems (NIDS) play a critical role in identifying unauthorized and malicious activities; however, conventional detection approaches often fail to cope with high-dimensional and dynamic network traffic, leading to limited detection accuracy and excessive false alarms. This study presents a comprehensive review of ensemble learning-based approaches for enhancing intrusion detection performance. By integrating multiple machine learning models through techniques such as bagging, boosting, and stacking, ensemble methods offer improved detection capability, robustness, and adaptability against both known and emerging attack patterns. An extensive analysis of existing research demonstrates that ensemble-based NIDS consistently outperform single-classifier models across widely used benchmark datasets, including NSL-KDD, UNSW-NB15, and CICIDS2017. Despite these advantages, issues related to computational overhead, class imbalance, and real-time deployment remain challenging. This review highlights current advancements, identifies open research challenges, and underscores the significance of dataset selection and ensemble design strategies in developing efficient and scalable intrusion detection solutions.

Keywords: Ensemble approach · Network Intrusion Detection Systems · Cybersecurity · Hyperparameters · Stacking

1 Introduction

Growing reliance on digital infrastructure has made network and system security a top issue. Intrusion Detection Systems enable surveillance and identification of malicious behavior as well as illegal access inside a network environment. Attacks on networks are increasing. Conventional intrusion detection systems find great challenges in the increasing complexity of modern network traffic, which increases false positive rates and reduces detection accuracy. Using ensemble learning techniques presents a workable solution for these challenges [1].

Ensemble techniques, such as bagging [2], boosting [3], and stacking [4], enhance the resilience, accuracy, and generalization capacity of models by integrating the strengths of various machine learning algorithms.

These strategies diminish the probability of mistakes associated with a singular model by amalgamating predictions from other models. According to recent research, while maintaining a low false positive rate, integrated intrusion detection systems can significantly improve the identification of both known and new types of attacks. In this paper, we explore the use of integrated methods in Network Intrusion Detection Systems (NIDS) to generate more effective and flexible intrusion detection mechanisms in dynamic and continuously growing network environments. This review paper analyzes the application of integration methods in network intrusion detection systems (NIDS), highlighting key datasets, methods, and their effectiveness in identifying various forms of network intrusions. The objective is to comprehensively analyze contemporary practices and highlight the performance superiority of integrated learning methods in the field of network intrusion detection systems (NIDS). Figure 1 represents the overview of the framework for Network Intrusion Detection Systems.

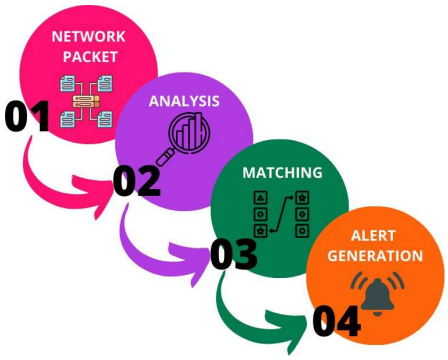

Fig. 1. General Design of Network Intrusion Detection Systems.

1.1 Key Contribution

- This article discusses the enhancement of Network Intrusion Detection Systems (NIDS) using ensemble learning approaches.
- The research explores numerous ensemble learning strategies and their efficacy in enhancing accuracy and resistance against both known and innovative network threats.
- The study evaluated the efficacy of ensemble methods on various datasets, including NSL-KDD, KDD-CUP99, UNSW-NB15, CICIDS2017, and CIDDS- 01.
- The research highlighted main problems in the implementation of ensemble techniques, including processing requirements, data imbalance, and system integration complications.

1.2 Article Organization

The article is structured as follows. Section 2 provides a literature review and an outline of important features and limitations. Section 3 presents the dataset utilised for network intrusion detection systems. Section 4 delineates the ensemble strategies, including their classifications, advantages, and limitations. The final part, 5, addresses the conclusion.

2 Related Work

Several studies have explored the effectiveness of ensemble techniques in Network Intrusion Detection Systems (NIDS). Alhowaide et al. [5] applied a Model Selection Method on datasets including NSL-KDD, UNSW-NB15, BoTNeTIoT, and BoTIoT, obtaining high F-scores in the range 0.95 to 1, so demonstrating a great detection capability across several invasions. M. Rashid et al. [6] proposed a tree-based stack ensemble model, obtained an accuracy of 99.9% on NSL-KDD and 94% on UNSW-NB15 datasets. Similarly, Stiawan et al. [7] used an ensemble approach on the ITD-UTM dataset, integrating different classifiers to generate accuracy rates between 81% and 85%. This draws attention to the variations in integrated method success rates depending on the employed datasets and approaches.

Thanh et al. [8] used six ensemble methods to identify DoS attacks on the UNSW-NB15 dataset, obtaining an F-measure of 99.28%, therefore proving the effectiveness of ensemble learning for certain attack types. Bukhari et al.[9] Implemented various ensemble models using UNSW-BC15 and CICIDS2017 datasets. The overall accuracy of the ensemble method is approximately 80.25%, while the boosting method reached 98.6%, and the stacking method achieved 98.8%. This indicates that the methods of boosting and stacking are more effective than other methods. Parashar et al. [10] presented a stacked ensemble framework that confirmed that this approach is useful in enhancing the performance of intrusion detection systems by delivering a 99% accuracy rate when evaluated on the CICIDS2017 dataset. With an accuracy of 99.68% on the NSL-KDD and UNSW-NB15 datasets. The IDS-EFS model by Akhiat performed excellently, therefore validating the efficacy of the mix of feature selection and ensemble approaches [11].

Furthermore, Alsaffar et al. [12] applied stacked ensemble learning on the UNSW-NB15 and CICIDS2017 datasets, achieving 99.92% accuracy on the latter, while performance on the former was lower at 84.88%, highlighting the variation in effectiveness across different datasets. Lastly, on the InSDN dataset, Hasanain et al. [12] developed a hybrid feature selection strategy and combined it with an ensemble soft voting classifier to attain almost perfect accuracy (99.9%). This research review reveals that, across various datasets and attack types, feature selection—especially through integrated approaches like stacking and boosting highly improves the accuracy of network intrusion detection systems (NIDS). The chosen dataset and the particular integration techniques used usually determine the success of these models. Table 1 presents the key features and challenges of various studies.

Table 1. Key Features and Challenges of Ensemble Techniques in NIDS.

Ref	Model Employed	Key Features	Challenges
[5]	Model Selection Method	Strong detection capability, robust performance across various intrusions	Performance variations based on method implementation
[6]	Tree-based stacking ensemble	High accuracy with complex attack detection	Potential difficulty in generalization to broader environments
[7]	SU, BN, OR, and J48 ensemble	Integration of multiple classifiers for flexibility in detection	Lower overall performance compared to other ensemble techniques
[8]	Six ensemble techniques for detecting DoS attacks	Effective for specific attack categories like DoS	Limited applicability to wider range of network attacks
[9]	Various ensemble models (boosting, stacking)	High efficiency with boosting and stacking methods	General ensemble performance lower than specialized methods
[10]	Stacking ensemble framework	High-level performance in improving IDS detection	Model evaluated on a specific scenario, limiting generalization
[11]	IDS-EFS	Strong feature selection improves detection rates	Model might face overfitting when applied to other environments
[12]	Stacked ensemble learning	Highly accurate method with strong generalization capability	Inconsistent performance across different scenarios
[13]	Hybrid feature selection with ensemble soft voting	Nearly perfect accuracy with improved feature selection	Dependent on specific features selected for optimal performance

3 Datasets for NIDS Evaluation

This section covers commonly available datasets with their characteristics and limitations.

3.1 Overview of the Common Dataset

In the domain of Network Intrusion Detection Systems (NIDS), a variety of publicly available datasets have been employed for testing and evaluation purposes. NSL-KDD [14], KDD-CUP99 [15], CIDDS-01 [16], CICIDS2017 [17], and UNSW-NB15 [18], as presented in Figure 2 are the commonly available datasets, exhibit significant variability in terms of class numbers and data counts, as well as differences in their collection methodologies [19]. Among these, the NSL-KDD dataset is the most often used one for NIDS. It was developed to produce a more refined dataset by removing duplicated or pointless entries, therefore addressing the flaws in the KDD-CUP99 dataset.

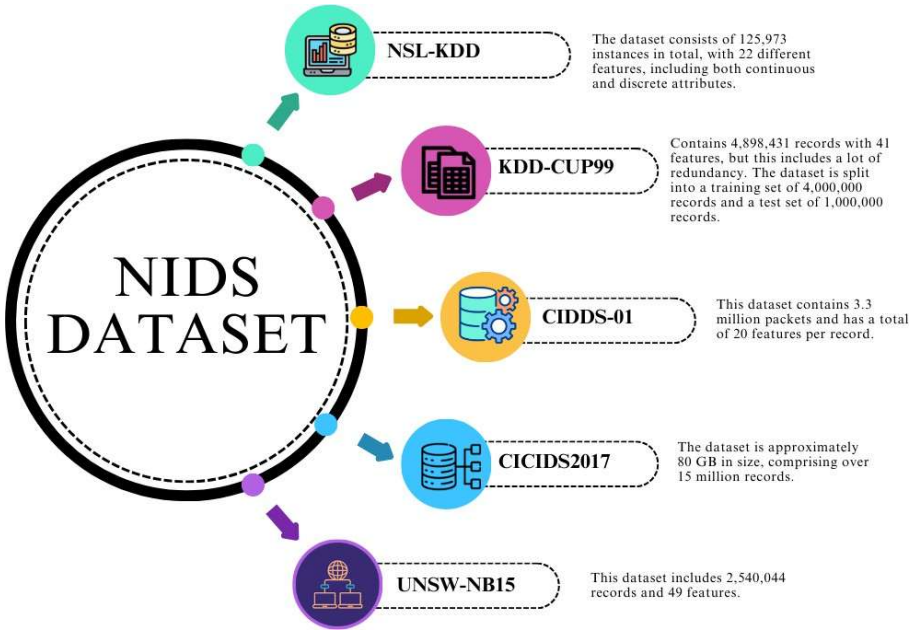

Fig. 2. Commonly available NIDS Dataset.

3.2 Dataset Characteristics and Limitations

Table 2 provides the key characteristics and limitations of each dataset.

4 Ensemble Techniques in NIDS

A review of ensemble learning, several ensemble approaches, and the benefits and drawbacks of these methods are given in this section.

4.1 Overview of Ensemble

Ensemble refers to the integration of outcomes from various learners to achieve dependable forecasts. This can be accomplished through many methodologies,

Table 2. Characteristics and Limitations of Common NIDS Datasets

Dataset	Characteristics	Limitations
NSL-KDD [14]	Contains 125,973 instances with balanced attack types. Reduces redundancy compared to KDD-CUP99.	Limited to older attack patterns, not reflective of modern threats. Does not capture real-world traffic accurately.
KDD-CUP99 [15]	Large dataset with over 4.8 million records, widely used as a benchmark.	High redundancy in data, leading to biased results. Lacks modern attack scenarios.
CIDDS-01 [16]	Focuses on real-world network traffic and includes both application and network layer attacks.	Limited dataset size and documentation, may not cover a wide range of attacks.
CICIDS2017 [17]	Large dataset with realistic network conditions, blending real and synthetic attack data.	Complexity and large size may hinder quick analysis and require significant computational resources.
UNSW-NB15 [18]	Reflects modern attack techniques with a detailed feature set. Suitable for contemporary security challenges.	Smaller number of attack types, and some simulated data may not represent real-world scenarios perfectly.

including the use of diverse datasets or learning frameworks. The primary problem in ensemble learning is selecting the algorithms and the decision or fusion function that integrates the outcomes of these algorithms. Dietterich [20] elucidates the application of ensemble-based systems for empirical validation, computational characterization, and representation.

4.2 Types of Ensemble Learning Methods

Ensemble construction entails the formulation and amalgamation of base classifiers, with three prevalent methodologies being bagging, boosting, and stacking. These strategies reduce discrepancies in predictions from bagging, boosting, or stacking [21]. Figure 3 displays the visual representation of these techniques.

Bagging Techniques Bagging is a simple ensemble approach that produces diverse outcomes by randomly sampling subsets from the training dataset. Various classifiers are developed utilising training data, and the ensemble's decision is dictated by the majority of classifiers for any given instance data [2]. Random Forests is a classifier that employs bagging, utilizing several decision trees with randomly varied parameters. It can generate training data and create unique subsets of attributes [22]. A new strategy called "pasting small votes" has arisen in ensemble learning as a variant of the established bagging method. This method is specifically designed for use with huge datasets, overcoming the constraints of conventional bagging [23].

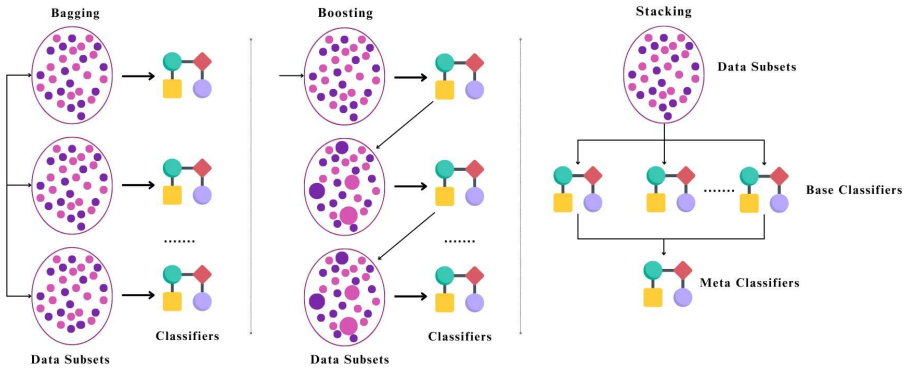

Fig. 3. Ensemble Designs Bagging, Boosting and Stacking.

Govindarajan et al. [24] introduce a distinctive ensemble classification method utilizing bagging classifiers, specifically Support Vector Machine (SVM) and Radial Basis Function (RBF), to enhance classification accuracy. The performance evaluation of NSL-KDD datasets indicates that the bagged RBF classifier achieved an accuracy of 86.40%, but the bagged SVM classifier reached a superior accuracy of 93.92%.

Boosting Techniques In 1990, Schapire [25] devised a weak learning method that generates classifiers outperforming random guessing. This approach, termed boosting, develops a strong learner by combining many classifiers via data resampling and majority vote. The boosting mechanism entails the development of three classifiers: the initial classifier is trained on a randomised subset of the training data; the second classifier employs an informed subset, comprising instances accurately classified by the first classifier alongside those misclassified; and the third classifier is trained on instances where the first two classifiers exhibit disagreement. The ultimate categorisation is established through a majority vote among these classifiers.

Freund and Schapire [26] developed this idea in 1997 by presenting "adaptive boosting," sometimes known as "AdaBoost," which has later been embraced generally in machine learning. Two well-known variants of AdaBoost that shine in multiclassification and regression are AdaBoostM1 and AdaBoostR. AdaBoost uses specific assumptions based on which weighted majority voting can help to consolidate classifier decisions. It uses an iterative approach whereby the distribution of the training data is changed to highlight cases misclassified by previous classifiers, therefore ensuring that next classifiers focus on ever more challenging examples. This adaptive approach increases classifier efficiency and general learning process.

Stacking Techniques In machine learning classification, misclassification frequently arises when cases are positioned close to the decision boundary defined by classifiers. This closeness may result in erroneous classifications, whereas clearly defined examples that are further from the boundary are more likely to be appropriately categorised. The study investigates the possibility of forming links between the outcomes of several classifiers applied to datasets from unspecified sources, with the objective of enhancing group detection accuracy.

Stacked Generalisation, commonly referred to as stacking, is presented as a technique to improve classification efficacy by utilising the outputs of a classifier ensemble. This method involves training a supplementary "meta-learner" to discern the correlations between the ensemble's predictions and the accurate labels, so enhancing the classification procedure. Stacking contrasts with conventional ensemble techniques like bagging and boosting, which generally emphasise the formation of homogeneous ensembles. Stacking focuses the integration of several learning environments, maybe resulting in better results on classification. This work emphasizes the need of stacking as a useful method to improve classifier performance on challenging datasets [4].

4.3 Advantages of Ensemble Techniques in NIDS

In network intrusion detection systems (NIDS), ensemble approaches provide many technological benefits. Often achieving identification accuracy of above 99% for network intrusions across many benchmark datasets, techniques such as random forests, gradient boosting, and stacking, which improve detection accuracy, reduce false positives, and handle imbalanced datasets [27].

4.4 Challenges in Implementing Ensemble Techniques

The adoption of Ensemble methods in Network Intrusion Detection Systems (NIDS) offers significant advantages, including improvements in accuracy and detection effectiveness. However, these models face many challenges. They are usually computationally demanding and can hinder real-time detection, especially when processing large datasets. Furthermore, data imbalance poses a significant challenge. This is because malicious traffic usually constitutes only a small portion of network data, making it more challenging to detect specific types of attacks. The visualization of ensemble models is more difficult than that of simple classifiers, further complicating the understanding and response to threats. The integration with existing systems and the response to ever-evolving attack patterns present additional challenges. Challenges such as hyperparameter tuning [6], overfitting [28], and managing large-scale networks complicate its application.

5 Conclusion

The rising complexity of network traffic and growing cyberattacks call for sophisticated solutions for network intrusion detection systems (NIDS). By combining many machine learning techniques—bagging, boosting, stacking—ensemble learning methods have shown great potential in improving the accuracy and durability of NIDS. Studies show that these techniques can improve the detection of both known and new attack types while reasonably lowering false positive rates.

Still difficult, though, are computing requirements, data imbalance, and integration with current systems. Notwithstanding these challenges, the use of ensemble approaches in NIDS marks significant progress in the continuous endeavor to protect digital infrastructures from developing vulnerabilities. The paper emphasizes the need to choose suitable datasets and integration techniques to maximize performance, therefore stressing the possible power of ensemble learning as a reliable method in the field of cybersecurity.

References

1. Hung-Jen Liao, Chun-Hung Richard Lin, Ying-Chih Lin, and Kuang-Yuan Tung. Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36(1):16–24, 2013.
2. Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.
3. Robert E Schapire et al. A brief introduction to boosting. In *Ijcai*, volume 99, pages 1401–1406. Citeseer, 1999.
4. David H. Wolpert. Stacked generalization. *Neural networks*, 5(2):241–259, 1992.
5. Alaa Alhowaide, Izzat Alsmadi, and Jian Tang. Ensemble detection model for iot ids. *Internet of Things*, 16:100435, 2021.
6. Mamunur Rashid, Joarder Kamruzzaman, Tasadduq Imam, Santoso Wibowo, and Steven Gordon. A tree-based stacking ensemble technique with feature selection for network intrusion detection. *Applied Intelligence*, 52(9):9768–9781, 2022.
7. Deris Stiawan, Ahmad Heryanto, Ali Bardadi, Dian Palupi Rini, Imam Much Ibnu Subroto, Mohd Yazid Bin Idris, Abdul Hanan Abdullah, Bedine Kerim, Rahmat Budiarto, et al. An approach for optimizing ensemble intrusion detection systems. *Ieee Access*, 9:6930–6947, 2020.
8. Hoang Ngoc Thanh and Tran Van Lang. Use the ensemble methods when detecting dos attacks in network intrusion detection systems. *EAI Endorsed Transactions on Context-aware Systems and Applications*, 6(19):e5–e5, 2019.
9. Owais Bukhari, Parul Agarwal, Deepika Koundal, and Sherin Zafar. Anomaly detection using ensemble techniques for boosting the security of intrusion detection system. *Procedia Computer Science*, 218:1003–1013, 2023.
10. Anshu Parashar, Kuljot Singh Saggu, and Anupam Garg. Machine learning based framework for network intrusion detection system using stacking ensemble technique. *Journal of Network and Computer Applications*, 201:102934, 2022.
11. Yassine Akhiat, Kaouthar Touchanti, Ahmed Zinedine, and Mohamed Chahhou. Ids-efs: Ensemble feature selection-based method for intrusion detection system. *Multimedia Tools and Applications*, 83(5):12917–12937, 2024.
12. Ali Mohammed Alsaffar, Mostafa Nouri-Baygi, and Hamed M Zolbanin. Shielding networks: enhancing intrusion detection with hybrid feature selection and stack ensemble learning. *Journal of Big Data*, 11(1):1–32, 2024.
13. Hasanain Ali Al Essa and Wesam S Bhaya. Ensemble learning classifiers hybrid feature selection for enhancing performance of intrusion detection system. *Bulletin of Electrical Engineering and Informatics*, 13(1):665–676, 2024.

14. Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A Ghorbani. Nsl-kdd dataset. https://www.unb.ca/cic/datasets/nsl.html, 2009. Accessed: 2024-10-04.
15. Salvatore J Stolfo, Wenke Fan, Wei Lee, et al. Kdd cup 99 dataset. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, 1999. Accessed: 2024-10-04.
16. Matthias Ring, Sascha Wunderlich, Dieter Landes, and Joachim Diederich. Cidds-001 dataset. https://www.hs-coburg.de/forschung/projekte-produktionen/cidds.html, 2017. Accessed: 2024-10-04.
17. Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani. Toward generating a new intrusion detection dataset and intrusion traffic characterization. https://www.unb.ca/cic/datasets/ids-2017.html, 2018. Accessed: 2024-10-04.
18. Nour Moustafa and Jill Slay. Unsw-nb15 dataset. https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/, 2015. Accessed: 2024-10-04.
19. Farida Suleiman, Umar Iliyasu, and Mukhtar Abubakar. Review on the network intrusion detection systems (nids). *BIMA JOURNAL OF SCIENCE AND TECHNOLOGY (2536-6041)*, 8(3A):141–155, 2024.
20. Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, 2000.
21. Majid Torabi, Nur Izura Udzir, Mohd Taufik Abdullah, and Razali Yaakob. A review on feature selection and ensemble techniques for intrusion detection system. *Faculty of Computer Science and Information Technology, Universiti Putra Malaysia*, 2022.
22. Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
23. Leo Breiman. Pasting small votes for classification in large databases and on-line. *Machine learning*, 36(1–2):85–103, 1999.
24. M. Govindarajan. Hybrid intrusion detection using ensemble of classification methods. *International Journal of Computer Networks & Information Security*, 6(2):45–53, 2014.
25. Robert E. Schapire. The strength of weak learnability. *Machine learning*, 5(2):197–227, 1990.
26. Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.
27. Sabrine Ennaji, Nabil El Akkad, and Khalid Haddouch. A powerful ensemble learning approach for improving network intrusion detection system (nids). In *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*, pages 1–6. IEEE, 2021.
28. Arindam Sarkar, Hanjabam Saratchandra Sharma, and Moirangthem Marjit Singh. A supervised machine learning-based solution for efficient network intrusion detection using ensemble learning based on hyperparameter optimization. *International Journal of Information Technology*, 15(1):423–434, 2023.