STM JOURNALS

International Journal of
Wireless Security and Networks

ISSN: 2584-0703
Volume 3, Issue 2, 2025
July–December
DOI (Journal): 10.37591/IJWSN

https://journals.stmjournals.com/ijwsn

Review IJWSN

DDoS Detection Using Cascade Correlation for Improving Network Resources in Cloud Environment

D. Prabakar[1,*], Naveena M.[2], Deepika S.[2], Pavithra K.[2]

Abstract

Intrusion detection is critical for protecting network security from emerging cyber threats. This study describes a unique intrusion detection system (IDS) based on the Random Forest algorithm. Random Forests are used as an effective classifier to identify patterns linked with malevolent behaviour. This technique uses Random Forests to improve the accuracy and efficiency of intrusion detection systems. The suggested methodology's value is shown by its performance on the benchmark KDD Cup dataset, which shows significant gains in detection accuracy when compared to previous methodologies. This study improves intrusion detection technology by demonstrating Random Forests' effectiveness in overcoming the hurdles presented by complex assaults. The system employs Random Forests as its primary classification method to accurately detect and analyze behavioral patterns linked to different types of cyberattacks, including denial-of-service (DoS), probing, remote-to-local (R2L), and user-to-root (U2R) intrusions. By generating numerous decision trees during the training phase and determining the final output based on the majority vote among these trees, Random Forests deliver improved performance, especially in managing noisy datasets and minimizing the risk of overfitting—issues that often challenge conventional intrusion detection methods.

Keywords: DDoS, intrusion detection, machine leaning, *random forests, cloud environment*

INTRODUCTION

In today's digital era, computer network and data security are critical. Given the growing complexity of cyber threats and the interconnectedness of our systems, effective network intrusion detection systems (NIDS) have never been more critical. Intrusion detection is critical for organizational security because it identifies illegal access and reduces possible dangers to information systems. Traditional intrusion detection technologies struggle to adapt to the dynamic threat environment. To address these issues and improve intrusion detection efficacy, we propose a novel technique named "Network Intrusion Detection with Two-Phased Hybrid Ensemble Learning and Automatic Feature Selection". This effort aims to combine cutting-edge approaches from machine learning, data science, and cyber security. Our goal is to improve network intrusion detection by combining the power of ensemble learning with automated feature selection in a two-phase detection strategy [1].

In the digital era, we are surrounded by tremendous volumes of information. Data has become the lifeblood of the information age, with site clicks, smartphone sensors, and vast databases powering modern organizations. However, in this data-driven era, the capacity to translate raw data into useable information is what lays the

***Author for Correspondence**
D. Prabakar
E-mail: prabakar.d@kce.ac.in

[1]Professor & Head of the Department, Department of Computer Science and Engineering, Karpagam College of Engineering Coimbatore, Tamil Nadu, India
[2]UG Scholar, Department of Computer Science and Engineering, Karpagam College of Engineering Coimbatore, Tamil Nadu, India

Received Date: March 24, 2025
Accepted Date: July 11, 2025
Published Date: September 09, 2025

Citation: D. Prabakar, Naveena M., Deepika S., Pavithra K. DDoS Detection Using Cascade Correlation for Improving Network Resources in Cloud Environment. International Journal of Wireless Security and Networks. 2025; 3(2): 17–22p.

groundwork for the most incredible technical developments we have ever seen. At the core of this revolutionary process is the science of "machine learning". Machine learning was driven by our need to make sense of the huge and complicated data ecosystems that comprise our modern world. It is motivated by the awareness that traditional rule-based programming is sometimes unable to handle the complexities of today's scenarios. Instead, machine learning allows computers to learn from data, discover patterns, and form conclusions, often with amazing precision. Machine learning algorithms may be trained to forecast future events or outcomes, such as stock prices, weather patterns, and medical diagnoses [2].

LITERATURE REVIEW

In the proposed technique, modern ransomware families use powerful encryption and dissemination techniques, making data recovery almost impossible. We look at the usage of software-defined networking (SDN) to identify and prevent complex ransomware attacks. We explain the findings from our ransomware investigation, as well as the SDN-based security architecture we created. The renowned Winery ransomware acted as a proof-of-concept. Based on our findings, we create an SDN detection and mitigation framework and provide an OpenFlow based solution. The proposed solution identifies suspicious activity via network traffic monitoring and blocks infected hosts by adding flow table entries to Open Flow switches in real time. Finally, our testing with many WannaCry samples demonstrate that the proposed system can quickly detect infected PCs and prevent WannaCry from spreading. Nowadays, ransomware poses a significant and rising threat to all sorts of users, from modest families to huge enterprises and government entities. Beginning with fairly basic phony antivirus software in 2008, five ransomwares developed over time, culminating in more complex versions such as crypto ransomware. The apotheosis of this evolution is the introduction of a new kind of ransomware that exploits flaws and spreads via worm-like ways across internal and external networks [3].

In their research, Benkeser *et al.* claimed that effective network intrusion detection solutions are crucial for addressing emerging cybersecurity threats. Historically, conventional business networks have received substantial attention in this subject. However, the cyber threat environment now includes wireless networks. The authors of this study provide a unique model that can be trained on entirely different feature sets and used to two independent intrusion detection applications: traditional corporate networks and 802.11 wireless networks. This is the first approach to demonstrate improved performance in both of the aforementioned applications. The model is built on a one-versus-all binary architecture with several stacked sub-ensembles. To achieve excellent generalization, each sub-ensemble has a collection of sub-learners, with just a subset of the sub-learners using boosting. Each class's sub ensembles are allocated a class weight based on the sensitivity measure (true-positive rate), which is determined entirely from training data. The use of pruning to remove sub-learners that do not contribute or have a negative impact on overall system performance is being researched [4].

Lei *et al.* suggested in their study that Intrusion detection systems (IDS) are widely used in network architecture to ensure the integrity and availability of sensitive assets in secure systems. Although numerous supervised and unsupervised machine learning methods have been tried to improve the efficacy of intrusion detection systems, existing algorithms continue to perform badly. First, high-dimensional datasets include a huge quantity of redundant and extraneous data, which disrupts an IDS's classification process. Second, an individual classifier may not be effective in detecting all types of assaults. Third, many models are built on outdated data, making them less adaptive to new threats. Thus, in this study, we provide a novel intrusion detection system based on feature selection and ensemble learning approaches. The first stage is to develop CFS-BA, a heuristic approach for dimensionality reduction that picks the best subset based on feature correlation. Then, we provide an ensemble technique that integrates the C4.5, Random Forest (RF), and Forest by Penalizing Attributes (Forest PA) algorithms [5].

Cheng *et al.* introduced a system as network security is critical for maintaining secure communication while minimizing financial loss and service disruptions caused by network attacks. Intruders often use holes in popular software to launch a range of assaults on networked computer systems. Network assaults may cause minor service disruptions to significant financial losses. Recently, intrusion detection systems (IDSs) based on machine learning techniques have evolved to tackle illegal network resource consumption and access. Over time, several machine learning approaches have been developed and used to IDSs. However, the vast majority of intrusion detection systems reported poor performance in terms of false positives and detection rates. To address these challenges, researchers developed ensemble classifiers, which integrate predictions from several separate classifiers. Ensembled classifiers compensate for the inadequacies of individual classifiers by pooling their expertise and improving performance [6].

According to Li *et al.*, in today's information era, a web-based attack protection system is needed. Classifier ensembles have been proposed for detecting anomaly-based intrusions in internet traffic. However, their performance is poor due to ineffective ensemble design. This study presents a stacked ensemble for anomaly-based intrusion detection systems in a web application. Unlike traditional stacking, which often employs single weak learners, the proposed stacked ensemble employs ensemble architecture with base learners such as random forest, gradient boosting machine, and XG Boost. To demonstrate the generalizability of the proposed method, the experiment employs two datasets designed particularly for Web application attack detection, CSIC-2010v2 and CICIDS-2017. Furthermore, the proposed model outperforms existing web attack detection algorithms in terms of accuracy and false positive rate. Validation results from the CICIDS 2017, NSL-KDD, and UNSW-NB15 datasets outperformed those obtained using several current approaches. Finally, the performance of all categorization algorithms is assessed using a two-step statistical significance test, which adds to the current literature [7].

RELATED WORK

Machine learning-based network intrusion detection systems (NIDSs) take use of flow characteristics produced by flow exporting protocols like NetFlow. ML and Deep Learning (DL) based NIDS systems, which have recently shown success, are anticipated to extract the average packet size as well as other flow information from each packet in the flow. In fact, flow exporters are often used on commodity devices when packet sampling is necessary. As a consequence, it is unclear if such machine learning based network intrusion detection systems are beneficial when sampling is available (flow information is obtained from a subset of packets rather than the whole traffic). In this work, we investigate how packet sampling affects the effectiveness and performance of machine learning based NIDSs. Unlike earlier research, our suggested assessment method is insensitive to changes in flow export stage characteristics. As a result, even with sampling, it can provide a reliable estimate of NIDS. We learned from sample studies that even at modest sampling rates such as 1/10 and 1/100, malicious flows of smaller size (i.e., packet count) are likely to escape unnoticed [8].

METHODOLOGY

The suggested system employs the Random Forest algorithm to detect intrusions in a dynamic and constantly changing cyber threat environment. The system establishes a solid basis by importing important data, including the well-known KDD dataset. Subsequent data pre-processing resolves concerns with cyber security data, ensuring that it is clean and ready for analysis. To enhance intrusion detection algorithms, feature selection divides characteristics into four categories: Basic, Content, Traffic, and Hosts. During the training and testing phases, the Random Forest technique is used to detect patterns and correlations in the data, and the model's efficacy is measured using important metrics such as Detection Rate (DR) and False Alarm Rate. The result of these modules is an Intrusion Detection System (IDS) that uses machine learning approaches, namely the Random Forest algorithm, to successfully improve cyber security measures against growing threats in the digital world [9].

Load Data

The initial stage in developing the suggested hybrid intrusion detection system (IDS) is loading the dataset. The dataset should include characteristics taken from network traffic. Several approaches may be used to extract the characteristics, such as packet sniffer, flow analysis, and network activity analysis [10].

Data Pre-processing

After importing the dataset, it must be pre-processed to improve the capabilities of an Intrusion Detection System (IDS). The module gives a comprehensive evaluation of the algorithm's performance in detecting cyber dangers in the Internet environment. The IDS's performance may be measured using a range of measures, including accuracy, precision, recall, and the F1 score. Accuracy refers to the percentage of network traffic that is accurately categorized. Precision refers to the fraction of malicious traffic that is successfully recognized. Recall refers to the proportion of malicious traffic identified, regardless of whether it was properly categorized. The F1 score represents the harmonic mean of accuracy and recall.

Computing

This necessitates cleaning up the data to eliminate any noise or outliers. It may also include transforming the data to a format compatible with the IDS's machine learning capabilities.

Feature Selection

Not all of the dataset's characteristics will be relevant for intrusion detection. Some features may be redundant or unnecessary. Feature selection is the process of determining which attributes are most essential for intrusion detection. This may be accomplished using a variety of techniques:

- Loading Dataset,
- Pre-processing,
- Feature Selection,
- Result,
- Anomaly Score,
- Based on Selected Features: Detecting Threats Using RF Method.

Training and Testing

Once the characteristics are determined, the IDS may be trained. This involves training machine learning algorithms to detect whether network traffic is legitimate or malicious. The trained models may now be used to categorize fresh network traffic. To evaluate the IDS's performance, it must be tested against a holdout set. The test set should include data not utilized to train the IDS. This enables for an unbiased evaluation of the IDS's performance.

Intrusion Detection Using Random Forest Algorithm

The last session focuses on the Random Forest technique to intrusion detection.

ALGORITHM DETAILS

Random Forest is a supervised learning approach. It is often used for classification and regression. How exactly does the Random Forest algorithm work? The Random Forest approach is broken into two steps: random forest creation and prediction using the random forest classifier created in the first phase. The author first shows the Random Forest building pseudocode:

1. Choose "K" features at random from a total of "m" features, where k is fewer than m.
2. Select the optimal split point to find the node "d" among the "K" attributes.
3. Split the node into daughter nodes using the optimum split.
4. Repeat steps a-c until the "l" number of nodes is reached.
5. Make a forest by repeating steps a-d "n" times to get "n" trees.

The following stage will be to create a prediction using the produced random forest classifier. The comparison chart is depicted in Figure 1.

1. Takes the test features, predicts the result using the rules of each randomly created decision tree, and records the anticipated outcome (target).
2. Total the votes for each anticipated objective.
3. The random forest algorithm's most voted projected target serves as the basis for the final forecast. The technique is straightforward, but also efficient.

RESULT ANALYSIS

The study's findings indicate that a Random Forest based intrusion detection system enhances the accuracy and efficiency of identifying hostile behaviours in network data. The suggested strategy outperforms earlier techniques in terms of detection accuracy, as shown by an assessment on the KDD Cup benchmark dataset. This implies that Random Forests may identify complex patterns associated with sophisticated attacks, showing their potential to strengthen network security measures against rising cyber threats. Overall, the findings highlight the practical benefits of adding Random Forest algorithms into intrusion detection systems to reduce risks and strengthen network defences (Table 1).

An intrusion detection system (IDS) that uses the Random Forest algorithm aims to improve accuracy and efficiency in detecting hostile activities in network data. The unique strategy is evaluated on the KDD Cup dataset, and it shows considerable gains in detection accuracy when compared to earlier techniques. In agreement with these findings, the provided results reveal that the Random Forest based IDS achieves an excellent 85% accuracy rate. This exceeds previous algorithms like Decision Trees (DT), which had an accuracy of 70%. The significant gap highlights the Random Forest algorithm's value in identifying complex patterns associated with sophisticated cyber-attacks, reaffirming its reputation as a viable tool for improving network security.

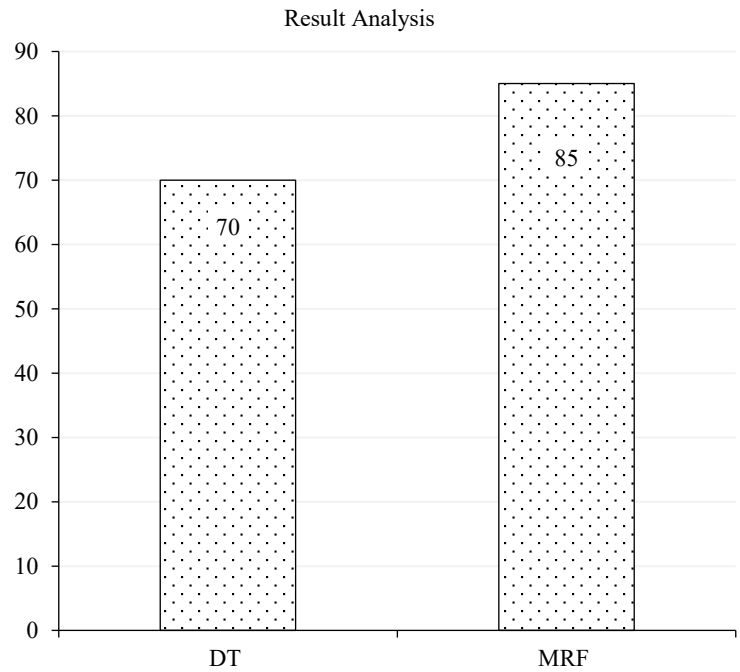

Figure 1. Result analysis of random forest.

Table 1. Comparison table.

Algorithm	Accuracy (%)
DT	70
MRF	85

CONCLUSION

Finally, the suggested intrusion detection system (IDS) based on the Random Forest algorithm offers a potential approach to improving network security against changing cyber threats. The IDS outperforms prior systems in terms of accuracy and efficiency since it employs Random Forests as a robust classifier. Evaluation on benchmark datasets demonstrates that the suggested method is effective in identifying patterns associated with harmful activities. This study advances intrusion detection technology by demonstrating the usefulness of Random Forests in overcoming the hurdles presented by sophisticated assaults. Moving ahead, further research and development in this area is required to investigate innovations and enhancements that will eventually boost network defences and protect important digital infrastructure.

Future Work

Future research in this field may focus on a range of exciting approaches to improving the efficacy and capabilities of intrusion detection systems that use the Random Forest algorithm. One alternative way is to look into ensemble learning methods, which integrate many classifiers, most notably Random Forests, to enhance detection accuracy and resilience. Furthermore, merging new anomaly detection approaches with deep learning frameworks might improve the IDS's capacity to detect novel and complex cyber threats.

REFERENCES

1. Kumar R, Malik A, Ranga V. An intellectual intrusion detection system using Hybrid Hunger Games Search and Remora Optimization Algorithm for IoT wireless networks. Knowl-Based Syst. 2022 Nov 28; 256: 109762.
2. Wang W, Jian S, Tan Y, Wu Q, Huang C. Representation learning-based network intrusion detection system by capturing explicit and implicit feature interactions. Comput Secur. 2022 Jan 1; 112: 102537.
3. Oughton EJ, Lehr W, Katsaros K, Selinis I, Bubley D, Kusuma J. Revisiting wireless internet connectivity: 5G vs Wi-Fi 6. Telecommun Policy. 2021 Jun 1; 45(5): 102127.
4. Benkeser D, Ju C, Lendle S, van der Laan M. Online cross-validation-based ensemble learning. Stat Med. 2018 Jan 30; 37(2): 249–60.
5. Lei S, Xia C, Li Z, Li X, Wang T. HNN: A novel model to study the intrusion detection based on multi-feature correlation and temporal-spatial analysis. IEEE Trans Netw Sci Eng. 2021 Sep 2; 8(4): 3257–74.
6. Cheng Y, Xu Y, Zhong H, Liu Y. Leveraging semisupervised hierarchical stacking temporal convolutional network for anomaly detection in IoT communication. IEEE Internet Things J. 2020 Jun 9; 8(1): 144–55.
7. Li X, Zhu M, Yang LT, Xu M, Ma Z, Zhong C, Li H, Xiang Y. Sustainable ensemble learning driving intrusion detection model. IEEE Trans Dependable Secure Comput. 2021 Mar 17; 18(4): 1591–604.
8. Alkanhel R, El-kenawy ES, Abdelhamid AA, Ibrahim A, Alohali MA, Abotaleb M, Khafaga DS. Network Intrusion Detection Based on Feature Selection and Hybrid Metaheuristic Optimization. Comput Mater Contin. 2023 Feb 1; 74(2): 2677–2693.
9. Kumar G, Thakur K, Ayyagari MR. MLEsIDSs: machine learning-based ensembles for intrusion detection systems—a review. J Supercomput. 2020 Nov; 76(11): 8938–71.
10. Tama BA, Nkenyereye L, Islam SR, Kwak KS. An enhanced anomaly detection in web traffic using a stack of classifier ensemble. IEEE Access. 2020 Feb 4; 8: 24120–34.

