A Two-Level Ensemble Learning Framework for Improved Accuracy and Reduced False Alarms in Network Intrusion Detection

Dr.K. Balamurugan[1], D. Navigneshkumar[2], Dr. G. Singarvel[3], Dr. S. Anuguraj[4], Mr.T. Sathish kumar[5], Mrs.S. Suganya[6]

[1]*Associate professor, Department of Information Technology, K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

[2]*M.Tech, Research Scholar, Department of Information Technology K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

[3]*Professor, Department of Information Technology, K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

[4]*Associate Professor, Department of Information Technology, K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

[5]*Assistant professor, Department of Information Technology, K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

[6]*Assistant professor, Department of Information Technology, K.S.R. College of Engineering (Autonomous), Tiruchengode-637 215, Namakkal, Tamilnadu, India*

Abstract- As cyber threats grow increasingly sophisticated, robust network security demands adaptive intrusion detection systems (IDS). Traditional machine learning-based IDS often struggle with high false alarm rates and poor generalization to emerging attacks, while deep learning-based IDS offer high detection accuracy but require significant computational resources. Ensemble learning techniques provide an effective balance between efficiency and accuracy, improving detection through model diversity and decision aggregation. This review explores ensemble-based intrusion detection systems, emphasizing diverse aggregation techniques, including homogeneous and heterogeneous ensemble methods. It provides an in-depth analysis of feature selection strategies, data balancing techniques, and classification models, offering a comparative assessment across benchmark datasets. Additionally, the study highlights key challenges and outlines futureresearch directions to advance ensemble learning in network intrusion detection.

Keywords: Ensemble Learning, NIDS, Feature Selection, Machine Learning, Data Imbalance.

I. INTRODUCTION

Intrusion Detection Systems (IDS) aim to detect unauthorized use, misuse, and abuse of computer networks by insiders and external attackers. Traditional IDS rely on the assumption that intruders' behavior will markedly differ from legitimate users, making unauthorized actions detectable. The advancement of Artificial Intelligence (AI) has spurred the development of fully automated IDS utilizing AI techniques such as neural networks support vector machines decision trees naive bayes and random forest. However, most AI methods, except decision forest, are base learning models that do not combine decisions for IDS management. These models often suffer from high false positive rates, with large companies receiving thousands of security alerts daily, and high false negative rates, which pose risks in safety-critical network applications. Previous studies typically focused on the classification accuracy of individual AI algorithms without harnessing the collective strength of these diverse techniques. This has highlighted the urgent need to utilize ensemble learning methods to enhance IDS effectiveness.

1.1. OVERVIEW

A wireless sensor network called the sound surveillance system (SOSUS) was created by the United States Military around sixty years ago. Its purpose was to detect and track submarines [1]. The technology back then was very cumbersome and pricey. Distributed sensor network (DSN) is the name given to the sensor technology in 1980. Scientists have investigated and resolved several problems and difficulties related to DSN [1], [2]. There have been enormous changes in this technology from that time to the 21st century. Sensor technology has become more pervasive in modern life. It combines the detected data of several physical world activities and turns it into useful information. Everything is carried out in a decentralized fashion. Wireless sensor networks (WSNs) can detect a wide variety of events, including changes in temperature and vibrations, the movement of objects, and even intrusion. The distance in geometric terms between the event and the sensor nodes determines the sensing capabilities. Small or large-scale WSNs may be formed by enumerable tiny wireless sensor nodes that are autonomous and powered by batteries. This technique is known as sensor technology.

1.2. Ensemble Learning for Binary Classification Anomaly Detection Approache

The article introduces an anomaly detection framework using datasets like CICIDS-2017, UNSW-NB15, and KDD'99, applying feature selection via the Chi-square method, and leveraging base models such as Gaussian Naive Bayes, Logistic Regression, and Decision Tree. Predictions are generated using a Stochastic Gradient Descent ensemble model. This study highlights the use of stacking to enhance IDS performance across various datasets, though it notes limitations such as data imbalance, suggesting that data augmentation could help in solving this imbalanace limitation.

II. LITERATURE SERVEY

2.1. Adaptive Ensemble Learning Framework for Robust Intrusion Detection in Wireless Sensor Networks

AUTHOR- VinodVeeramachaneni,Corresponding Author

YEAR-2025

DESCRIPTION

In recent years, the security of Wireless Sensor Networks (WSNs) has faced significant challenges due to their susceptibility to various cyber threats. This study introduces an Adaptive Ensemble Learning Framework that enhances the robustness of intrusion detection in WSNs. The framework leverages a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, integrated through a dynamic voting mechanism to improve detection accuracy and reduce false positives.

2.2. A Two-Level Ensemble Learning Framework for Enhancing Network Intrusion Detection Systems

AUTHOR- OSVALDO ARRECHE1, ISMAIL BIBERS2, and MUSTAFA ABDALLAH2

YEAR-2023

DESCRIPTION

The exponential growth of intrusions on networked systems inspires new research directions on developing artificial intelligence (AI) techniques for intrusion detection systems (IDS). In this context, several AI techniques have been leveraged for automating network intrusion detection tasks. However, each AI model has unique strengths points and weaknesses, and one may be better than the other depending on the dataset, which might aggravate which model to choose.

2.3. Ensemble Learning-Based Intrusion Detection and Classification for Securing IoT Networks: An Optimized Strategy for Threat Detection and Prevention

AUTHOR- Kumaresh Sheelavant1, Charan K. V. 2, B. Yamini Supriya3, Purshottam J. Assudani4, Chandra Bhushan Mahato5, Sanjay Kumar Suman6

YEAR-2025

DESCRIPTION

The Internet of Things (IoT) advancement has created new security holes, which require intrusion detection systems to defend networks effectively. The complex

structure of IoT networks causes traditional security methods to fail because they produce high amounts of incorrect detections and limited ability to accurately identify threats. The authors introduce ID-ELC: Ensemble Learning and Classification framework for Intrusion Detection, which aims to strengthen IoT environment security.

2.4. Ensemble-IDS: An Ensemble Learning Framework for Enhancing AI-Based Network Intrusion Detection Tasks

AUTHOR- Ismail Bibers, Osvaldo Arreche

YEAR-2025

DESCRIPTION

Modern cybersecurity threats continue to evolve in both complexity and prevalence, demanding advanced solutions for intrusion detection. Traditional AI-based detection systems face significant challenges in model selection, as performance varies considerably across different network environments and attack scenarios.

2.5. Improving Threat Detection in Information Security with Ensemble Learning

AUTHOR- Ahmad Sanmorino, RendraGustriansyah, ShintaPuspasar

YEAR-2025

DESCRIPTION

In the face of increasingly complex and frequent cyberattacks, traditional rule-based threat detection systems often fail to identify evolving malicious behaviours. This study addresses the challenge by leveraging ensemble learning to enhance intrusion detection in information security. By integrating three distinct machine learning models – Support Vector Machine (SVM), Random Forest, and Deep Neural Network (DNN) – the proposed approach capitalises on their strengths while mitigating their weaknesses. The primary goal is to enhance detection accuracy, minimise false positives, and ensure reliable performance across various attack types. Using benchmark datasets

III. PROJECT DESCRIPTION

3.1. EXISTING SYSTEM

The growth of Internet and the services provided by it has been growing exponentially in the past few decades. With such growth, there is also an ever-increasing threat to the security of networks. Several efficient countermeasures have been placed to deal with these threats in the network, such as the intrusion detection system (IDS). This paper proposes an ensemble learning-based method for building an intrusion detection model. The model proposed in this paper has relatively better overall performance than its individual classifiers. This ensemble model is constructed using lightweight machine learning models, i.e., Gaussian naive Bayes, logistic regression and decision tree as the base classifier and stochastic gradient descent as the meta-classifier.

3.1.1. DISADVANTAGE

- The increasing number of users and their dependence on the Internet, there has been an increase of attacks in the network that disrupts its normal functioning.
- Many different types and scales of attacks are on the rise. Security policy, antiviruses and firewalls are no longer enough to protect the network.
- A system that is designed to protect against targeted attacks by constantly monitoring traffic of the network is the network intrusion detection system (NIDS). Such systems analyze the data and try's to identify any anomaly in the network such as unauthorized access, alteration, damage or intrusions.

3.2. PROPOSED SYSTEM

The scale of the development of Internet of Things (IoT) contributed to the level of vulnerability to cyber-attacks, causing severe disruption of services and identity theft in the case of unidentified attacks. In order to increase the security of the IoT, an IDS based on deep learning was created that used a four-layer Full Connected network to detect malicious traffic via IoT devices. The system was communication protocol-independent and it made deployment much easier and found 93.74 % attacks including Black hole, DDoS, Sinkhole, and Wormhole. The experimental analysis revealed that the performance was strong, and the average precision and recall, and F1-scores are more

than 93, which guarantees reliability of the IoT network protection.

Network attacks were a significant issue to the contemporary digital systems, and they impacted all networks regardless of their size. A system of detection of such threats and its mitigation required an IDS. Machine learning and deep learning technologies were extensively used to develop effective security systems, but it was necessary to develop more flexible solutions due to the changes in malicious actions.

3.2.1. ADVANTAGE

- A hybrid deep-learning-based intrusion detection system (HDLNIDS) was created with the assistance of a convolutional recurrent neural network, where CNN layers learned local features, and RNN layers learned sequential patterns.
- The model was tested on the CICIDS-2018 dataset and had 98.90% accuracy, which is superior to the current detection methods.

3.3. MODULE DESCRIPTION

1. Dataset Collection and Preprocessing

This study is based on the fact that a credible and representative dataset is used which captures a broad range of network behaviors, both benign and malicious. In this respect, the main data source is selected as NSL-KDD dataset. It is a more polished version of the KDD99 dataset having been carefully edited to eliminate duplicated records, bias, and provide a more balanced sample distribution. NSL-KDD contains labeled network traffic examples, which are classified into normal or different types of intrusions such as DoS, Probe, R2L, and U2R. The dataset consists of 41 features that describe the features of network connections like duration, type of protocol, service, flag, and other statistical measures taken when making a connection.

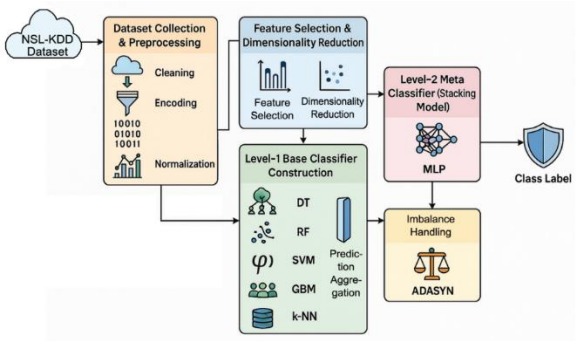
2.Feature Selection and Dimensionality Reduction

Since the preprocessing is already done, the next thing is to narrow down the dataset by isolating the most informative features. This is critical in alleviating the curse of dimensionality, speeds up training time, and the overall learner's generalization capability. The technique of mutual information (MI) is used to perform feature selection. MI measures the information that is exchanged between any two features of input and the target variable.

3. Prediction Aggregation

After training and validation of the base classifiers, their outputs are collected to create a new dataset which is commonly known as the meta-feature matrix. With a given sample, each model in its base form produces the expression of a discrete prediction or a probability distribution of the possible predictions. These outputs are never considered as final predictions but rather they are used as inputs in a higher order learning model.

3.4.BLOCK DIAGRAM



IV. RESULT AND DISCUSSION

Intrusion detection systems (IDSs) are an essential element for network security infrastructure and play a very important role in detecting large number of attacks. They are widely used systems to detect malicious intent activities and various attacks on the internet. Although there are different types of IDS, all these systems suffer from a common problem which is generating high volume of alerts and huge number of false alarms. This lessens the proficiency of the IDS. This drawback has become the main motivation for many research papers in IDS area. The aim of conducted research in the field is to propose different techniques to handle the alerts, reduce them and distinguish real attacks from false positives and low importance events. This paper is a compilation of research in this field which reviews existing false alarm minimization techniques in signature-based Network Intrusion Detection System (NIDS).

V. CONCLUSION

A framework for Software-Defined Networking (SDN) forensic readiness and cyber security incident response integrates the standard NIST incident response lifecycle with SDN-specific considerations, using a block diagram that shows the phases of Preparation, Detection and Analysis, Containment, Eradication, and Recovery, and Post-Incident Activity. The diagram would connect these phases in a cycle, with sub-blocks for each phase detailing SDN-specific tasks, such as preparing the SDN controller for forensic logging, identifying threats through SDN flow and traffic data, isolating compromised network segments via the controller, and analyzing forensic evidence from SDN components in the post-incident phase.

VI. FUTURE ENHANCEMENTS

The expanding Internet of Things operational space requires IDS systems to protect IoT devices because their rapid growth produces security risks. Different detection and mitigation methods have been developed for cyber-attacks through mechanisms that preserve both low resource usage and high accuracy rates. The detection methods extend from statistical evaluation to machine learning when combined with hybrid models. The survey evaluates multiple intrusion detection methodologies as well as their strengths and weaknesses for IoT systems. The selection process for features stands as a vital component for improving the operational effectiveness of intrusion detection models. PCA and Fisher Dimension Reduction methods serve to eliminate superfluous features thus achieving better efficiency with no impact on detection success rates. Mutual information-based approaches have been utilized to find the most crucial attributes for achieving classification purposes. The performance-improving techniques show limitations when detecting different attack patterns due to their reduced accuracy level.

REFERENCE

[1] K. Wang, A. Zhang, H. Sun, and B. Wang, “Analysis of recent deep learning-based intrusion detection methods for in-vehicle network,” IEEE Trans. Intell. Transp. Syst., vol. 24, no. 2, pp. 1843–1854, Feb. 2023.

[2] A. Prasanth and S. Jayachitra, “A novel multi-objective optimization strategy for enhancing quality of service in IoT-enabled WSN applications,” Peer-to-Peer Networking and Applications, vol. 13, no. 6, pp. 1905-1920, 2020.

[3] A. Smith, B. Johnson, and C. Lee, “A Machine Learning Approach for Cyber Threat Detection in IoT Environments,” Journal of Network and Computer Applications, vol. 175, pp. 102900, Jan. 2023.

[4] C. Kalaiselvi and G. M. Nasira, “A new approach for diagnosis of diabetes and prediction of cancer using ANFIS,” in Proc. World Congress on Computing and Communication Technologies, WCCCT 2014, pp. 188-190, 2014.

[5] M. Thiruvengadam et al., “Bioactive compounds in oxidative stress-mediated diseases: Targeting the nrf2/are signaling pathway and epigenetic regulation,” Antioxidants, vol. 10, no. 12, pp. 1-12, 2021.

[6] M. Zakariah, S. A. AlQahtani, and M. S. Al-Rakhami, “Machine learning-based adaptive synthetic sampling technique for intrusion detection,” Applied Sciences, vol. 13, no. 11, p. 6504, 2023.

[7] V. D. P. Jasti et al., “Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis,” Security and Communication Networks, vol. 2022, pp. 1-12, 2022.

[8] V. Roy, L. Roy, R. Ahluwalia, G. Khambra, M. Ramesh, and K. Rajasekhar, “An advance implementation of machine learning techniques for the prediction of cervical cancer,” in Proc. 3rd IEEE Int. Conf. ICT in Business Industry and Government, ICTBIG 2023, 2023.

[9] S. K. Suman et al., “Game theoretical approach for improving throughput capacity in wireless ad hoc networks,” in Proc. Int. Conf. Recent Trends in Information Technology, ICRTIT 2014, pp. 10-12, 2014.

[10] H. K. Shakya et al., “Energy-proficient cluster enrichment in wireless sensor networks via categorized fuzzy clustering and multi-hop routing optimization,” SN Comput. Sci., vol. 6, no. 25, 2025.