

Aode-Based Intrusion Detection with Subset Feature Selection Using an Optimized Cascade Correlation Neural Network

Uma S¹, Abi M², Anu Krithika P³, Anubharathi T⁴

¹Assistant Professor, Department of Information Technology, Panimalar Engineering College

^{2,3,4}IVyr Students, Department of Information Technology, Panimalar Engineering College

ABSTRACT

This research introduces an Intrusion Detection System (IDS) that uses the Anonymised Online Data Expansion (AODE) classification approach to improve network threat detection. The system, written in Java and including a graphical user interface, analyses network traffic data in ARFF format for training and assessment purposes. The AODE-based model enhances anomaly detection by analysing network traffic patterns with a probabilistic learning method. It effectively identifies known and new assaults, supports both real-time and offline modes, and adapts to changing threats. The system's user-friendly interface enables the visualisation and understanding of findings, assisting security administrators in automated decision-making. Its scalable architecture allows it to fit into a large variety of network environments, ranging from business to cloud infrastructures. By integrating with current security frameworks, the IDS strengthens cybersecurity defences, seeking to decrease false positives while retaining high detection rates.

Keywords: Electricity load stability, Machine learning models, Dimensionality reduction, Energy management optimization.

1. INTRODUCTION

As cybersecurity threats continue to evolve, having efficient Intrusion Detection Systems (IDS) is becoming more and more important to protect networks from malicious attacks. Conventional IDS models often struggle with accuracy and adaptability, particularly when detecting novel and emerging threats. This study introduces an IDS based on the Anonymised Online Data Expansion (AODE) classification algorithm, leveraging a probabilistic machine learning approach to enhance detection accuracy by incorporating attribute correlations. The system, which was implemented in Java, has a graphical user interface that is easy to use and inspects network traffic data in the form of ARFF to detect known and unknown intrusions. It supports both real-time and offline analysis, ensuring adaptability across diverse network environments, including enterprise and cloud infrastructures. Designed to integrate seamlessly with existing security frameworks, the IDS enhances defense mechanisms by reducing false positives while maintaining a high detection rate. By employing advanced machine learning techniques, this system provides an automated and scalable solution for proactive cybersecurity, aiding organizations in securing their digital assets against emerging threats.

ANONYMIZED ONLINE DATA EXPANSION (AODE)

Anonymised Online Data Expansion (AODE) is a machine learning classification approach for increasing the accuracy of probabilistic models. This classifier is an extension of the Naïve Bayes, specifically developed for complicated datasets with interdependent characteristics. AODE addresses all potential attribute relationships rather than addressing them independently, resulting in more accurate categorisation. This approach is very valuable in cybersecurity for intrusion detection since it may detect hidden patterns in network data. Unlike typical classifiers, AODE improves learning by dynamically increasing the training cases. It detects both known and undiscovered cyber threats accurately and efficiently. The approach effectively reduces false positives, making it a trustworthy option for IDS applications. Intrusion detection systems that use AODE can achieve greater accuracy and flexibility to emerging security threats.

ANOMALY-BASED DETECTION

Anomaly-based detection is particularly helpful for detecting zero-day attacks, distributed denial-of-service (DDoS) attacks, and other uncommon security threats that signature-based detection might overlook. This method does not require prior knowledge of specific attack signatures, making it useful for detecting previously identified threats. The system learns the baseline or "normal" behaviour from prior data and continually analyses incoming traffic or activities for deviations from it. An anomaly, such as a rapid rise in network traffic or an unexpected pattern of requests, is identified as a possible security issue. Anomaly-based detection is especially beneficial for identifying zero-day attacks, distributed denial-of-service (DDoS) assaults, and other unique security breaches that signature-based approaches may not detect. While it may provide false positives owing to innocuous fluctuations in behaviour, the approach is nonetheless an important component of proactive cybersecurity tactics.

2.LITERATURE REVIEW

2.1. Title: AN INTRUSION DETECTION SYSTEM FOR SWARMING OF DRONES USING TIMED PROBABILISTIC AUTOMATA

Author: Venkatraman Subbarayalu et al.

Year: 2024 This research tackles the security issues brought about by drone swarms, which are being used more and more in almost all fields including military operations, surveillance, and monitoring the environment. The authors suggest an intrusion detection system (IDS) utilizing Timed Probabilistic Automata (TPA) to emulate typical drone swarm behavior and identify anomalies that could signal an attack which could influence the efficiency and precision of the project. The proposed IDS is highly adaptive and capable of identifying various types of drone swarm intrusions, including zero-day attacks. Its flexibility and efficiency make it a valuable tool in safeguarding UAV networks from malicious infiltration.

2.2. Title: PERFORMANCE EVALUATION AND COMPREHENSIVE ENUMERATION OF 500+ NATURE-INSPIRED METAHEURISTIC OPTIMIZATION ALGORITHMS

Author: Zhongqiang Ma et al.

Year: 2024 This study presents a comprehensive performance assessment and categorization of over 500 nature-inspired metaheuristic algorithms. The authors analyze the increasing number of newly developed metaheuristics that draw inspiration from biological, physical, and human behavioral phenomena. Many of these algorithms are benchmarked against traditional optimization techniques without rigorous comparative evaluation on complex datasets. The study systematically classifies these algorithms based

on their source of inspiration and key solution-generating mechanisms, providing valuable insights for researchers in the field of optimization and algorithm development.

2.3. Title: TESTING AND OPTIMIZATION METHODOLOGIES ON STANDARD BENCHMARK FUNCTIONS OF LOAD FREQUENCY CONTROL IN INTERCONNECTED MULTI-AREA POWER SYSTEMS IN SMART GRIDS

Author: Krishan Arora et al.

Year: 2023 This research focuses on improving load frequency control in smart grid systems by integrating advanced optimization methodologies. The study proposes the Harris Hawks Optimizer (HHO) as an efficient solution for addressing frequency regulation challenges in interconnected multi-area power systems. By applying the HHO algorithm to various benchmark functions, including unimodal, multimodal, and fixed-dimension test cases, the researchers demonstrate its superior performance compared to existing optimization techniques. The results highlight the significance of optimization in ensuring stability and efficiency in modern smart grids powered by renewable energy sources.

2.4. Title: A RELIABLE HYBRID MACHINE LEARNING MODEL FOR NETWORK INTRUSION DETECTION

Author: Md. Alamin Talukder et al.

Year: 2022 This research proposes a new hybrid intrusion detection system (IDS) that uses machine learning and deep learning approaches to improve network security. The suggested model enhances anomaly-based intrusion detection using Synthetic Minority Over-sampling Technique (SMOTE) to balance the data and XGBoost for selecting features. The research compares the performance of the hybrid approach with conventional machine learning and deep learning models. The outcomes prove enhanced detection accuracy and reliability, and the suggested approach is a feasible solution for managing large-scale network security threats.

2.5. Title: A METHOD OF DDOS ATTACK DETECTION USING NATURAL SELECTION OF MODELS AND FEATURES

Author: Ruikui Ma et al.

Year: 2020 This research introduces a Distributed Denial of Service (DDoS) detection framework that leverages an optimized feature and model selection approach (FAMS) to enhance cybersecurity defenses. The framework consists of four stages: data preprocessing, feature extraction, model selection, and attack detection, ensuring a structured approach to threat identification. The authors analyze 79 extracted features and apply multiple selection techniques such as backward elimination, mutual information, Lasso.L1, and random forest to enhance detection accuracy and computational efficiency. Extensive experimentation on real-world datasets demonstrates that the proposed FAMS-based approach significantly improves generalization capability and prediction efficiency. In addition, the framework efficiently reduces false positives and is able to evolve with changing attack patterns, thus constituting a scalable and resilient solution for countering contemporary DDoS attacks in dynamic networks.

3. EXISTING SYSTEM

The existing intrusion detection systems (IDS) Fig 3.1 are severely hampered in handling network traffic analysis as a result of escalating complexity and magnitude of cyber attacks. Conventional IDS approaches usually find it difficult to identify suspicious behavior accurately, particularly in the case of unbalanced network data. They can produce false alarms or miss advanced attack patterns, leading to their decreased effectiveness. The industry is in ever-increasing demand for an ever-more adaptive and effective IDS to

properly detect anomalies and address security threats in real-time. Furthermore, typical IDS solutions make use of static rule-based or signature-based detection methods, making them weaker when dealing with zero-day attacks as well as other emerging threats. This results in compromised threat detection and response time and further weakens cybersecurity defenses.

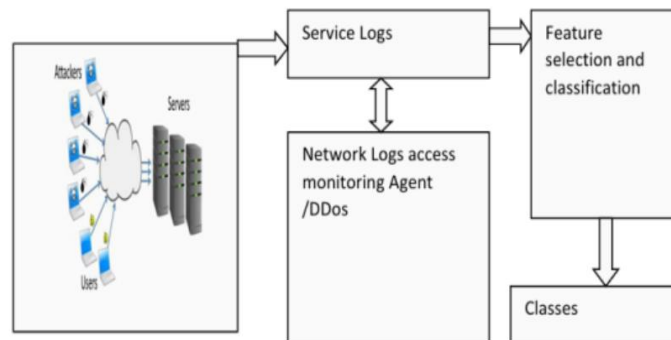


Fig 3.1 Existing System

The proposed system integrates a hybrid method of data preprocessing, involving oversampling the minority classes and undersampling the majority classes. The method balances network traffic data and avoids over-emphasis by the IDS on major patterns, making it better able to detect infrequent or unknown intrusions. Further, the Random Forest Regressor is utilized for attribute selection and is an essential component in identifying the most critical features that assist in intrusion detection. By using these primary attributes in model training, the system improves detection performance while minimizing computational complexity. The work makes an input in the enhancement of cybersecurity protocols and securing vital network infrastructures against ongoing threats.

4. PROPOSED SYSTEM

The proposed system Fig 4.1 introduces an AODE-based Intrusion Detection System (IDS) that enhances the accuracy of threat detection using machine learning techniques. Anonymized Online Data Expansion (AODE) classification method, the system effectively analyzes network traffic to identify and mitigate security threats. Developed in Java, the system has a friendly graphical user interface (GUI) to ease user control and interaction. It handles network datasets in ARFF format to effectively train and test the detection model. With probabilistic learning, the IDS can detect both known and previously unknown attacks with improved accuracy. Additionally, it supports real-time monitoring and offline analysis, allowing organizations to adapt the system based on their security requirements.

4.1 DATA PREPROCESSING MODULE

The data preprocessing module is an important component that manages and pre-processes network traffic data for analysis. It is tasked with transforming raw network logs into a well-structured format, such as ARFF, that can be used by machine learning algorithms. The preprocessing step includes data cleaning, where noisy and incomplete records are removed, normalization to scale numerical attributes, and filtering to eliminate redundant or irrelevant information. These processes ensure that the dataset is optimized for accurate and efficient model training.

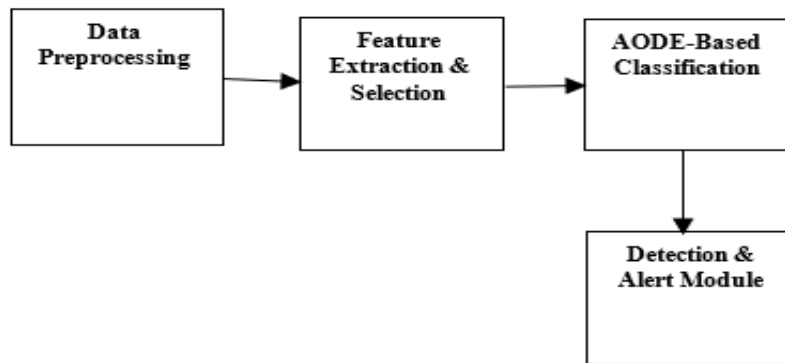


Fig 4.1 System Architecture

4.2 FEATURE EXTRACTION AND SELECTION MODULE

The feature selection and extraction module is concerned with the determination of the most critical attributes in the network dataset that are responsible for intrusion detection. This step involves analyzing network traffic patterns and isolating key indicators of potential threats to enhance the system's predictive capabilities. Various feature selection techniques are applied to prioritize critical features while eliminating less significant ones, thereby improving model efficiency. By reducing computational complexity, this module not only accelerates the training process but also ensures that the IDS focuses on the most meaningful data for precise threat identification.

4.3 AODE-BASED CLASSIFICATION MODULE

The AODE-based classification module, which is at the center of the proposed IDS, utilizes probabilistic learning methods to distinguish between malicious and normal network behaviors. The algorithm analyzes network traffic patterns and correlations between attributes to classify potential intrusions with a high level of accuracy. Unlike traditional IDS models, which often struggle with evolving threats, AODE's probabilistic approach enables it to detect both previously identified and emerging attack patterns. This module ensures that security breaches are identified efficiently while minimizing false positives, making it a reliable component for modern cybersecurity defense systems.

4.4 DETECTION AND ALERT MODULE

The detection and alert module is responsible for monitoring classified network traffic and generating timely notifications in response to potential threats. When an intrusion is detected, the system provides administrators with critical details, including the type of attack, severity, and timestamps for further analysis. This module supports real-time threat detection, allowing security teams to take immediate action in mitigating risks. Additionally, all detected threats are logged for documentation and future reference, enabling organizations to analyze trends in security breaches and refine their defensive strategies accordingly.

4.5 GRAPHICAL USER INTERFACE (GUI) MODULE

The GUI module improves usability by offering a graphical interface where the IDS may be managed by the user interactively. It allows users to load network datasets, configure system parameters, and visualize detection results through an intuitive dashboard. Security administrators can monitor network traffic patterns, review intrusion alerts, and adjust system settings based on organizational security needs. The user-friendly interface simplifies the operation of the IDS, making it accessible to both technical and non-technical users while ensuring effective cybersecurity management.

This proposed system offers a sample output fig 4.2 which shows advanced, efficient, and scalable solution

to intrusion detection, providing organizations with an automated defense mechanism against evolving cyber threats.

SAMPLE OUTPUT

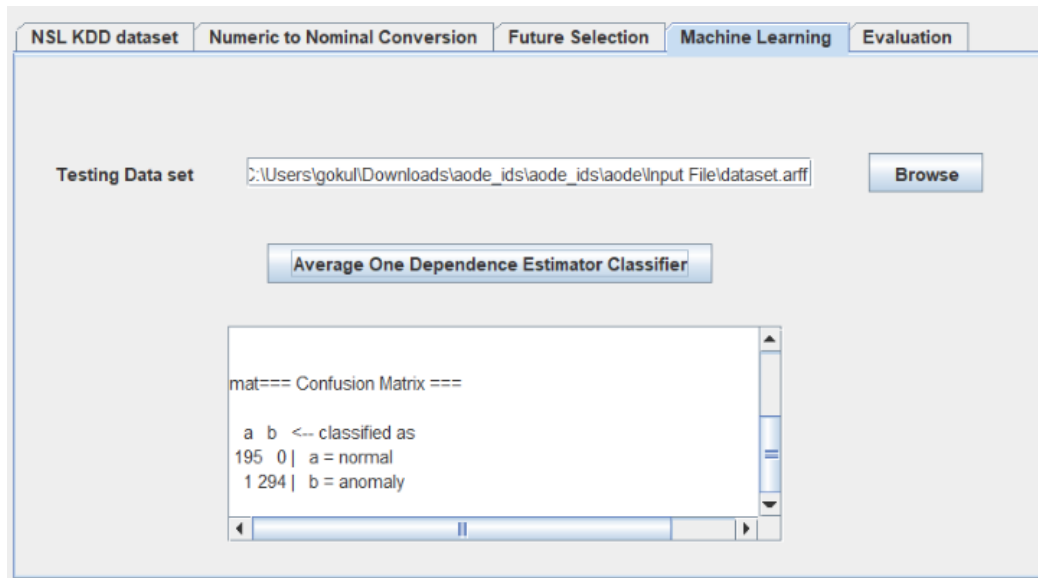
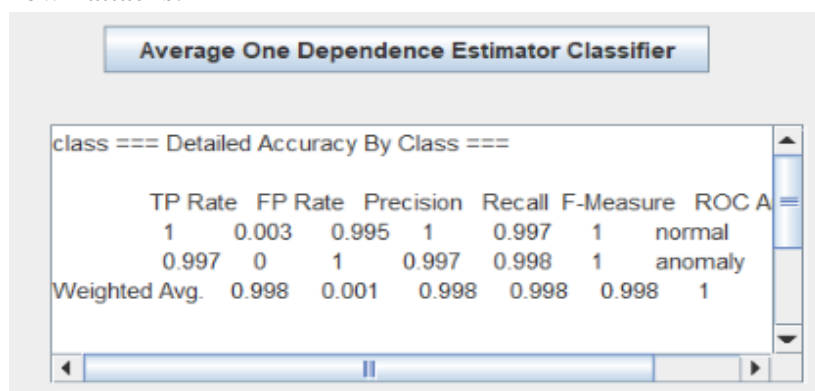


Fig 4.2 Sample Output

5. DISCUSSION & RESULT

The result Fig 5.1 is analysis of the AODE-based Intrusion Detection System (IDS) evaluates its performance in detecting malicious network activities. The system is tested using a benchmark dataset, where detection accuracy, false positive rate, and processing efficiency are analyzed. The AODE classification technique enhances accuracy by considering attribute dependencies, leading to improved anomaly detection. Comparison with conventional IDS models points out its strengths in the detection of both known and unknown attacks.



class === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC A
1	0.997	0.003	0.995	1	0.997	1
0	0.997	0	1	0.997	0.998	1
Weighted Avg.	0.998	0.001	0.998	0.998	0.998	1

Fig 5.1 Result

The effectiveness of the system is tested in terms of performance parameters like precision, recall, and F1-score. The detection outcome in real time is checked for promptness and reliability of the identification of intrusion. The IDS exhibits minimal false alarms at a high rate of detection efficiency. The IDS demonstrates a low false alarm rate while maintaining high detection efficiency. Graphical representations such as confusion matrices and detection trend charts help visualize system performance. Results indicate that AODE-based classification enhances intrusion detection capabilities over conventional methods. This

analysis validates the system's reliability, making it a valuable tool for cybersecurity defense.

6. CONCLUSION

The AODE-based Intrusion Detection System (IDS) offers a powerful and precise method of identifying cyber threats within network environments. Utilizing the Anonymized Online Data Expansion (AODE) classification method, the system improves the detection of known and unknown attacks. The structured input processing and feature extraction ensure reliable and efficient intrusion identification. The system's real-time monitoring and alert generation help administrators respond to threats promptly. Performance analysis demonstrates high accuracy, reduced false positives, and improved detection rates compared to traditional methods. The user-friendly graphical interface allows for easy interaction and result interpretation. Overall, this IDS contributes to strengthening network security and protecting digital assets from cyber threats. Future enhancements can further improve its adaptability to evolving attack patterns.

7. FUTURE SCOPE

Future work on the AODE-based Intrusion Detection System (IDS) will focus on enhancing its adaptability to emerging cyber threats. Implementing deep learning techniques can improve detection accuracy by identifying complex attack patterns. The system can be extended to support more real-time network environments with optimized processing speed. Integration with cloud-based security frameworks will enable scalable and distributed threat monitoring. Enhancing feature selection methods using advanced algorithms can further reduce false positives and improve efficiency. The system could incorporate self-learning mechanisms to adapt to new attack variations without manual updates. Adding a more detailed reporting and visualization module will improve user interaction and decision-making. Support for multiple data formats and real-time packet analysis will make the system more versatile. Mobile or web-based access to IDS reports and alerts can enhance usability for security teams. These future improvements will ensure the system remains robust and effective against evolving cybersecurity challenges.

8. REFERENCES

1. R. Kumar, A. Malik, and V. Ranga, "An intellectual intrusion detection system using hybrid hunger games search and remora optimization algorithm for IoT wireless networks," *Knowledge-Based Systems*, vol. 256, Nov. 2022, Art. no. 109762.
2. W. Wang, S. Jian, Y. Tan, Q. Wu, and C. Huang, "Representation learning-based network intrusion detection system by capturing explicit and implicit feature interactions," *Computers & Security*, vol. 112, Jan. 2022, Art. no. 102537.
3. J. Oughton, W. Lehr, K. Katsaros, I. Selinis, D. Bublely, and J. Kusuma, "Revisiting wireless internet connectivity: 5G vs Wi-Fi 6," *Telecommunications Policy*, vol. 45, no. 5, Jun. 2021, Art. no. 102127.
4. B. A. Tama and S. Lim, "Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation," *Computer Science Review*, vol. 39, Feb. 2021, Art. no. 100357.
5. S. Lei, C. Xia, Z. Li, X. Li, and T. Wang, "HNN: A novel model to study the intrusion detection based on multi-feature correlation and temporal spatial analysis," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 4, Oct. 2021, pp. 3257–3274.
6. A. K. Sarica and P. Angin, "A novel SDN dataset for intrusion detection in IoT networks," presented

at the 16th International Conference on Network and Service Management (CNSM), Izmir, Turkey, Nov. 2020, pp. 1–5.

7. S. Choudhary and N. Kesswani, "Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT," *Procedia Computer Science*, vol. 167, Jan. 2020, pp. 1561–1573.
8. P. Sun, P. Liu, Q. Li, C. Liu, X. Lu, R. Hao, and J. Chen, "DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system," *Security and Communication Networks*, vol. 2020, Aug. 2020, pp. 1–11, doi: 10.1155/2020/8890306.
9. N. Moustafa and A. Jolfaei, "Autonomous detection of malicious events using machine learning models in drone networks," presented at the 2nd ACM MobiCom Workshop on Drone-Assisted Wireless Communications for 5G and Beyond, Sep. 2020, pp. 61–66.
10. A. Akbar, M. Shameem, S. Mahmood, A. Alsanad, and A. Gumaei, "Prioritization-based taxonomy of cloud-based outsource software development challenges: Fuzzy AHP analysis," *Applied Soft Computing*, vol. 95, Oct. 2020, Art.no.106557, doi: 10.1016/j.asoc.2020.106557.