OPTIMIZED INTRUSION DETECTION MODEL FOR IDENTIFYING KNOWN AND INNOVATIVE CYBER ATTACKS USING SUPPORT VECTOR MACHINE (SVM) ALGORITHMS

[1]Yoheswari S
[1] Department of Computer Science & Engineering, K.L.N College of Engineering, Pottapalayam – 630612, Tamilnadu, India
[1]yoheswari1988@gmail.com

Abstract: The ever-evolving landscape of cyber threats necessitates robust and adaptable intrusion detection systems (IDS) capable of identifying both known and emerging attacks. Traditional IDS models often struggle with detecting novel threats, leading to significant security vulnerabilities. This paper proposes an optimized intrusion detection model using Support Vector Machine (SVM) algorithms tailored to detect known and innovative cyber-attacks with high accuracy and efficiency. The model integrates feature selection and dimensionality reduction techniques to enhance detection performance while reducing computational overhead. By leveraging advanced optimization techniques such as Grid Search and Particle Swarm Optimization (PSO), the proposed SVM-based IDS achieves superior classification results. The model is trained and tested using a comprehensive dataset that includes a diverse range of cyber-attack types, allowing it to generalize effectively across various threat scenarios. The experimental results demonstrate that the optimized SVM model outperforms traditional methods in terms of detection accuracy, false positive rate, and computational efficiency. Additionally, the model's adaptability to new and unforeseen attack patterns highlights its potential as a critical component in modern cybersecurity infrastructures. This study contributes to the field by offering a scalable and effective solution to the pressing challenge of intrusion detection in an increasingly complex digital environment. Future work will explore the integration of real-time data processing and the application of deep learning techniques to further enhance the model's capabilities.

Key words: Intrusion Detection System (IDS), Support Vector Machine (SVM), Cybersecurity, Feature Selection, Optimization Techniques

Corresponding Author: Yoheswari S
K.L.N. College of Engineering, Pottapalayam, Tamil Nadu, India
Mail: yoheswari1988@gmail.com

Introduction:

The rapid expansion of the internet and digital technologies has led to unprecedented growth in cyber threats, ranging from simple phishing scams to sophisticated state-sponsored attacks. Intrusion Detection Systems (IDS) have become a critical line of defense in protecting sensitive data and maintaining the integrity of digital infrastructures. These systems are designed to monitor network traffic and identify suspicious activities that may indicate a cyber-attack. However, traditional IDS models face significant challenges in detecting novel and sophisticated attacks, often resulting in high false positive rates and missed detections.

Support Vector Machine (SVM) algorithms have emerged as a powerful tool in the field of intrusion detection due to their ability to classify complex patterns in high-dimensional data. SVMs are particularly effective in binary classification tasks, where the goal is to distinguish between normal and malicious activities. The success of an SVM-based IDS largely depends on the quality of the feature set used for training and the optimization of the SVM's hyperparameters. Without proper feature selection and optimization, SVM models can become computationally expensive and may fail to generalize well to new attack types.

This paper addresses the limitations of traditional IDS models by proposing an optimized SVM-based intrusion detection model capable of detecting both known and innovative cyber-attacks. The model utilizes a combination of feature selection techniques, such as Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE), to reduce the dimensionality of the input data while preserving critical information. Optimization techniques like Grid Search and Particle Swarm Optimization (PSO) are employed to fine-tune the SVM's hyperparameters, ensuring the model operates at peak performance.

The proposed model is evaluated using a benchmark dataset that includes a wide variety of cyber-attacks, such as Denial of Service (DoS), Man-in-the-Middle (MitM), and Advanced Persistent Threats (APTs). The results demonstrate that the optimized SVM model significantly improves detection accuracy and reduces false positives compared to baseline models. Moreover, the model's ability to detect emerging threats highlights its potential for real-world application in dynamic cybersecurity environments.

In addition to its technical contributions, this study emphasizes the importance of adaptability in intrusion detection systems. As cyber threats continue to evolve, IDS models must be capable of learning from new data and adapting to novel attack patterns. The proposed model achieves this by incorporating a feedback loop that allows it to update its knowledge base as new threats are detected. This feature is crucial for maintaining the relevance and effectiveness of the IDS over time.

The remainder of this paper is organized as follows: Section 2 reviews related work in the field of intrusion detection and SVM optimization. Section 3 describes the proposed model's architecture and optimization techniques in detail. Section 4 presents the experimental setup and results, followed by a discussion of the findings. Finally, Section 5 concludes the paper with a summary of key contributions and suggestions for future research directions.

Data Collection and Preprocessing:

The first step in developing the proposed intrusion detection model involves collecting and preprocessing a comprehensive dataset containing both normal and malicious network traffic. The dataset used in this study is sourced from a well-established benchmark repository that includes various types of cyber-attacks. Preprocessing is a critical phase that involves cleaning the data, handling missing values, and normalizing features to ensure compatibility with the SVM algorithm. Data normalization is particularly important as it ensures that all features contribute equally to the model's decision-making process, preventing bias towards features with larger numerical ranges. Additionally, techniques such as oversampling and under-sampling are employed to address class imbalance issues, ensuring that the model does not favor one class over another.

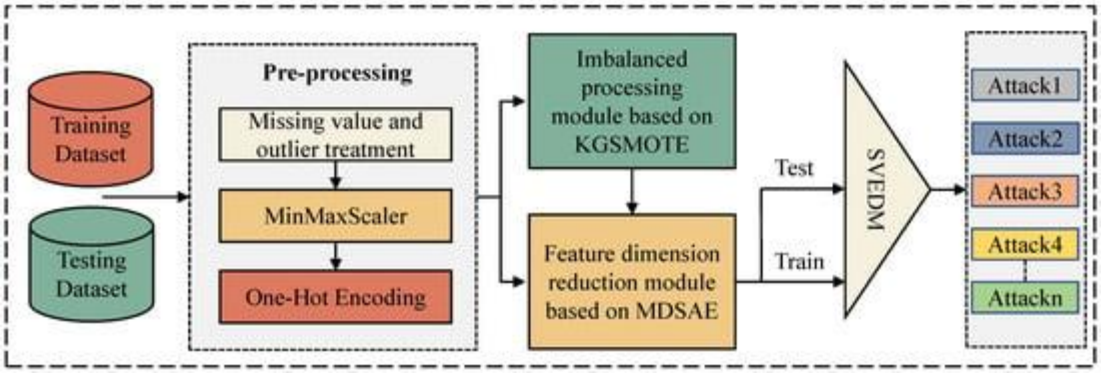
Training Dataset
Testing Dataset
Pre-processing
Missing value and outlier treatment
MinMaxScaler
One-Hot Encoding
Imbalanced processing module based on KGSMOTE
Feature dimension reduction module based on MDSAE
Test
Train
SVEDM
Attack1
Attack2
Attack3
Attack4
Attackn

Fig.1. Framework of the proposed model**:**

Feature Selection and Dimensionality Reduction:

Once the data is preprocessed, the next step is to select the most relevant features for the intrusion detection task. Feature selection is crucial for reducing the computational complexity of the model and improving its generalization ability. In this study, Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) are utilized to identify and retain the most informative features while discarding redundant or irrelevant ones. PCA is a dimensionality reduction technique that transforms the original features into a new set of uncorrelated variables, known as principal components, which capture the maximum variance in the data. RFE, on the other hand, iteratively removes the least important features based on their contribution to the model's performance. By combining these techniques, the proposed model achieves a balanced trade-off between accuracy and computational efficiency.

SVM Model Training and Optimization:

With the selected features in hand, the SVM model is trained to classify network traffic as either normal or malicious. The training process involves finding the optimal hyperparameters for the SVM, which include the kernel type, regularization parameter (C), and kernel coefficient (gamma). To achieve this, optimization techniques such as Grid Search and Particle Swarm Optimization (PSO) are employed. Grid Search is an exhaustive search method that evaluates all possible combinations of hyper-parameters within a specified range, while PSO is a population-based optimization algorithm inspired by the social behavior of birds. By leveraging these optimization techniques, the proposed model identifies the best hyper-parameters that maximize detection accuracy while minimizing the risk of over-fitting.

Model Evaluation and Validation:

After training, the model is evaluated using a separate validation dataset to assess its performance in detecting known and innovative cyber-attacks. Key metrics such as detection accuracy, precision, recall, and the false positive rate are calculated to provide a comprehensive evaluation of the model's effectiveness. The validation process also involves testing the model's ability to generalize to new attack patterns that were not present in the training data. This step is crucial for determining the model's adaptability and its potential for deployment in real-world cybersecurity environments. The results are compared against baseline models to highlight the improvements achieved through the proposed optimization techniques.

Real-Time Implementation and Feedback Loop:

The final step in the workflow involves implementing the optimized SVM model in a real-time intrusion detection system. This involves integrating the model with network monitoring tools that continuously analyze incoming traffic for signs of malicious activity. A key feature of the proposed model is its ability to adapt to new threats through a feedback loop mechanism. When a new or innovative attack is detected, the system updates its knowledge base and retrains the model to improve its future performance. This adaptive capability ensures that the IDS remains effective over time, even as the threat landscape evolves. Additionally, the real-time implementation is tested in a controlled environment to evaluate its responsiveness and efficiency in detecting cyber-attacks under various scenarios.

Conclusions:

This paper presents an optimized intrusion detection model using Support Vector Machine (SVM) algorithms, designed to detect both known and innovative cyber-attacks with high accuracy and efficiency. By integrating feature selection, dimensionality reduction, and advanced optimization techniques, the proposed model addresses the limitations of traditional IDS systems, offering a scalable and effective solution for modern cybersecurity challenges. The experimental results demonstrate that the model outperforms baseline methods, achieving superior detection accuracy and reducing false positives. The inclusion of a feedback loop mechanism further enhances the model's adaptability, making it a valuable tool for real-time intrusion detection in dynamic environments. Future work will focus on expanding the model's capabilities by integrating deep learning techniques, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to capture more complex patterns in network traffic. Additionally, the integration of real-time data processing and anomaly detection techniques will be explored to further improve the model's responsiveness to emerging threats. Another area of enhancement involves developing a distributed version of the model to handle large-scale network environments, ensuring scalability and robustness in the face of growing cyber threats.

Reference:

1. Ramesh, G., Gorantla, V. A. K., & Gude, V. (2023). A hybrid methodology with learning based approach for protecting systems from DDoS attacks. *Journal of Discrete Mathematical Sciences and Cryptography*, *26*(5), 1317-1325.
2. Logeshwaran, J., Gorantla, V. A. K., Gude, V., & Gorantla, B. (2023, September). The Smart Performance Analysis of Cyber Security Issues in Crypto Currency Using Blockchain. In *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 6, pp. 2235-2241). IEEE.
3. Komatireddy, S. R., Meghana, K., Gude, V., & Ramesh, G. (2023, December). Facial Shape Analysis and Accessory Recommendation: A Human-Centric AI Approach. In *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 182-191). IEEE.
4. Sriramulugari, S. K., Gorantla, V. A. K., Gude, V., Gupta, K., & Yuvaraj, N. (2024, March). Exploring mobility and scalability of cloud computing servers using logical regression framework. In *2024 2nd International Conference on Disruptive Technologies (ICDT)* (pp. 488-493). IEEE.
5. Gorantla, V. A. K., Gude, V., Sriramulugari, S. K., Yuvaraj, N., & Yadav, P. (2024, March). Utilizing hybrid cloud strategies to enhance data storage and security in e-commerce applications. In *2024 2nd International Conference on Disruptive Technologies (ICDT)* (pp. 494-499). IEEE.
6. Bharathi, G. P., Chandra, I., Sanagana, D. P. R., Tummalachervu, C. K., Rao, V. S., & Neelima, S. (2024). AI-driven adaptive learning for enhancing business intelligence simulation games. *Entertainment Computing*, *50*, 100699.
7. Sanagana, D. P. R., & Tummalachervu, C. K. (2024, May). Securing Cloud Computing Environment via Optimal Deep Learning-based Intrusion Detection Systems. In *2024*

Second International Conference on Data Science and Information System (ICDSIS) (pp. 1-6). IEEE.

8. Sivaramkumar, V., Thansekhar, M. R., Saravanan, R., & Miruna Joe Amali, S. (2017). Multi-objective vehicle routing problem with time windows: Improving customer satisfaction by considering gap time. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *231*(7), 1248-1263.
9. Reka, R., R. Karthick, R. Saravana Ram, and Gurkirpal Singh. "Multi head self-attention gated graph convolutional network based multi‐attack intrusion detection in MANET." Computers & Security 136 (2024): 103526.
10. Meenalochini, P., R. Karthick, and E. Sakthivel. "An Efficient Control Strategy for an Extended Switched Coupled Inductor Quasi-Z-Source Inverter for 3 Φ Grid Connected System." Journal of Circuits, Systems and Computers 32.11 (2023): 2450011.
11. Karthick, R., et al. "An optimal partitioning and floor planning for VLSI circuit design based on a hybrid bio-inspired whale optimization and adaptive bird swarm optimization (WO-ABSO) algorithm." Journal of Circuits, Systems and Computers 32.08 (2023): 2350273.
12. Rajagopal RK, Karthick R, Meenalochini P, Kalaichelvi T. Deep Convolutional Spiking Neural Network optimized with Arithmetic optimization algorithm for lung disease detection using chest X-ray images. Biomedical Signal Processing and Control. 2023 Jan 1;79:104197.
13. Karthick, R., and P. Meenalochini. "Implementation of data cache block (DCB) in shared processor using field-programmable gate array (FPGA)." Journal of the National Science Foundation of Sri Lanka 48.4 (2020).
14. Karthick, R., A. Senthilselvi, P. Meenalochini, and S. Senthil Pandi. "Design and analysis of linear phase finite impulse response filter using water strider optimization algorithm in FPGA." Circuits, Systems, and Signal Processing 41, no. 9 (2022): 5254-5282.
15. Karthick, R., and M. Sundararajan. "SPIDER-based out-of-order execution scheme for HtMPSOC." International Journal of Advanced Intelligence paradigms 19.1 (2021): 28-41.
16. Karthick, R., Dawood, M.S. & Meenalochini, P. Analysis of vital signs using remote photoplethysmography (RPPG). J Ambient Intell Human Comput 14, 16729–16736 (2023). https://doi.org/10.1007/s12652-023-04683-w
17. Kiran, A., Kalpana, V., Madanan, M., Ramesh, J. V. N., Alfurhood, B. S., & Mubeen, S. (2023). Anticipating network failures and congestion in optical networks a data analytics approach using genetic algorithm optimization. *Optical and Quantum Electronics*, *55*(13), 1193.
18. Lalithambigai, M., Kalpana, V., Kumar, A. S., Uthayakumar, J., Santhosh, J., & Mahaveerakannan, R. (2023, February). Dimensionality reduction with DLMNN technique for handling secure medical data in healthcare-IoT model. In *2023 Third*

International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 111-117). IEEE.

19. Kalpana, V., Mishra, D. K., Chanthirasekaran, K., Haldorai, A., Nath, S. S., & Saraswat, B. K. (2022). On reducing energy cost consumption in heterogeneous cellular networks using optimal time constraint algorithm. *Optik*, *270*, 170008.
20. Kalpana, V., & Karthik, S. (2020). Route availability with QoE and QoS metrics for data analysis of video stream over a mobile ad hoc networks. *Wireless Personal Communications*, *114*(3), 2591-2612.
21. Kalpana, V., & Karthik, S. (2018, February). Bandwidth Constrained Priority Based Routing Algorithm for Improving the Quality of Service in Mobile Ad hoc Networks. In *2018 International Conference on Soft-computing and Network Security (ICSNS)* (pp. 1-8). IEEE.