IRO Journal on Sustainable Wireless Systems (ISSN: 2582-3167)
www.irojournals.com/irosws/

Hybrid Feature Selection with Adaptive Threshold Optimization for Efficient Machine Learning–based Intrusion Detection

Valli S P[1], Rifat A K[2], Shameem Sakinah[3]

[1,2,3]Department of Computer Science and Engineering, B.S. Abdur Rahman Crescent Institute of Science and Technology, Chennai, India.

E-mail: [1]vallisp@crescent.education, [2]rifatkabeer805@gmail.com, [3]shameem.sakinah@gmail.com

Abstract

With the increasing use of computer networks and internet services, network security becomes crucially important. The existing intrusion detection methods are based on signature analysis or rules, but these methods are inefficient against zero-day attacks and evolving cyber threats. In order to overcome the disadvantages of existing solutions, this study proposes a novel IDS framework with hybrid feature selection method and adaptive threshold optimization. The IDS under consideration uses the UNSW-NB15 dataset. First, the data pre-processing algorithms such as encoding, scaling and class balancing are used. Hybrid feature selection is then implemented by means of a combination of chi-squared filtering and Particle Swarm Optimization (PSO). Several supervised learning algorithms, including K-Nearest Neighbors, Decision Tree, Logistic Regression and Random Forest are used for training and testing. Moreover, Adaptive Threshold Testing Algorithm (ATTA) is applied for dynamic optimization of decision thresholds. The results show that this solution considerably increases IDS efficiency and decreases false positive rates.

Keywords: Intrusion Detection System, Machine Learning, Feature Selection, Particle Swarm Optimization, Adaptive Threshold.

1. Introduction

The continuous evolution in computer networking, cloud computing, IoT, and internet service development has introduced drastic changes in the digital infrastructure. The

DOI: https://doi.org/10.36548/jsws.2026.3.005
Extended version of article received from **6th International Conference on Artificial Intelligence, 6G Communications and Network Technologies 2026**
Received:12.05.2026, received in revised form:16.06.2026, accepted: 03.07.2026, published: 14.07.2026
© 2026 Inventive Research Organization. This is an open access article under the Creative Commons Attribution-NonCommercial International (CC BY-NC 4.0) License

organizations are increasingly adopting such interconnected systems for their applications in various sectors including banking, health care, electronic commerce, and industrial automation. Nevertheless, this evolution has equally led to an increase in security challenges such that the networks have become vulnerable to cyber-attacks such as denial of service, malware spreading, privilege escalation, and data breach. Intrusion Detection Systems (IDS) help in monitoring and detecting any form of malicious activities, but in order to do that effectively, they must be adaptable, precise, and efficient in handling evolving attacks.

While ML-based IDS has improved the ability of detecting attacks by learning the behavior of traffic, however, it has to face some challenges including high dimensional datasets, feature redundancies, class imbalances, fixed decision threshold, etc. The high dimensionality causes overfitting, computational issues and imbalance on one hand, while on the other hand, high false positive rates accompany class imbalances. These false alarms reduce the system reliability and make it hard to implement in real-time large-scale networks.

Thus, feature selection and optimization of the decision thresholds become essential for IDS improvement. Feature selection helps in reducing the spatial dimension by overcoming the problem of feature redundancies and irrelevancies, thus improving the efficiency of the model. Using a combination of filter-based feature selection method such as chi-square selection and wrapper-based optimization technique such as Particle Swarm Optimization (PSO) ensures computational efficiency as well as improved classification performance. Moreover, adaptive threshold adjustment improves classification boundaries in order to improve precision and recall and reduce false positive rate, thus retaining good detection capabilities at the same time.

1.1 Existing System

The Intrusion Detection Systems (IDS) play a crucial role in ensuring the security of the emergence of modern computer networks with regard to the growing problem of cyberattacks. The development of the cloud computing technology and the growth of network traffic speeds and diverse devices made it necessary to employ rule-based and signature-based systems of attack detection in the past; however, they are not sufficient nowadays in detecting complicated and changing attacks. Despite the fact that machine learning algorithms applied in early IDS systems made it possible to increase the level of detection accuracy, such systems

were still ineffective in identifying zero-day threats, being unable to adapt to dynamic changes in traffic and detect non-linear relationships between different network features. Traditional feature extraction and selection techniques had an important disadvantage in that they inevitably lost significant discriminatory information. In order to deal with the aforementioned issues, feature selection methods utilizing optimization and hybrid machine learning models were developed to increase the accuracy and reduce the training time. Nevertheless, most of the approaches relied on fixed probability thresholds, thus reducing their flexibility in dealing with imbalanced datasets and changing network conditions. The lightweight IDS systems designed for resource-limited environments demonstrated lower computational overhead but had a limited experience in terms of testing on enterprise networks. Hybrid models which combined optimization algorithms with artificial neural networks demonstrated good results in terms of attack detection capability but still face the problems of high computational complexity, hardware dependency, and lack of interpretability which made them unfit for real-time systems. Finally, application layer IDS systems limited the scope of the attacks by omitting packet analysis and flow analysis.

More advanced techniques such as context-aware systems, resampling methods, scalable architectures, deep learning, multimodal systems, and reinforcement learning based IDS have been developed in order to enhance the adaptability of IDS. Some challenges related to additional computation cost, long inference time, unstable training processes, dependency on data from various data sources having heterogeneous data, as well as poor scalability for high dimensional data were associated with such systems. Most of the approaches based on deep learning technology are characterized by being a black-box approach, which has poor transparency and explainability for security analysts. In general, even though great advances have been achieved within the area of IDS research, certain limitations are evident in the aspect of adaptability, scalability, computational cost, interpretability and real-time implementation.

2. Related Works

Various researchers have worked towards IDS improvement by integrating advanced feature selection and machine learning technologies. An anomaly-based IDS that was introduced by Emirmahmutoglu and Atay [1] involved metaheuristic feature selection methods that include Particle Swarm Optimization (PSO) and Differential Evolution, and Flower

Pollination Algorithm. Their work showed better detection accuracy and less training time due to optimized feature subsets. Nevertheless, this system was based on fixed classification thresholds restricting the ability to adapt to changing network conditions. Kumar et al. [2] designed SG-IDS-a lightweight IDS for wireless sensor networks based on PCA and SVD dimension reduction, that uses classifiers such as Gaussian Naive Bayes and SGD to ensure efficient resource management. The algorithm was restricted to WSNs datasets and did not get tested at scale or enterprise. The hybrid IDS proposed by Bakır et al. [3] used Gravitational Search Algorithm (GSA) to optimize feature subset and Quantum Neural Network (QNN) to classify the attacks. Its performance was excellent owing to the feature selection and high learning capability. However, the algorithm could not be implemented in real-time due to its computational complexity and poor interpretability.

A phishing-based IDS that uses ensemble feature selection and neural network classifiers was proposed by Pattawaro et al. [4]. The IDS improved robustness and reduced overfitting when detecting web-based attacks. However, this approach applies only to application layer attacks and cannot detect network layer attacks. Turukmane et al. [5] proposed situational-based predictive IDS, a model of the network state as dynamic transitions between states. The IDS improved the multi-stage attack detection ability using contextual learning. However, feature selection is not used in the approach, leading to high computation costs in large data sets. Kareem et al. [6] proposed MLSTL-WSN using SMOTE-Tomek resampling to handle class imbalance in IDS data sets. The approach improved detection accuracy, especially for minority attack classes. However, the approach was tested in WSN environment only.

The scalable IDS framework by Alkanhel et al. [7] uses oversampling and feature embedding techniques. This IDS was highly accurate and improved the recall for detecting uncommon attacks. However, this IDS relied on static thresholds and it consumed more memory due to oversampling. To detect the spatial and temporal pattern of the attacks, Ahmed et al. [8] devised a hybrid IDS using XGBoost, CNN, and LSTM. The IDS is highly accurate for binary and multiclass classification problems. However, the computation of the model is highly expensive and has latency which limits the application of the IDS in real time. The multimodal IDS framework of Megantara et al. [9] involves network traffic, system logs and user behavior data. The combination of PCA and SVM is helpful in improving the accuracy

and reducing the false positive rate. However, the use of multiple data sources makes the implementation difficult.

A machine learning-based IDS for IoT networks was developed by Velliangiri et al. [10]. It improved the detection performance of resource-constrained IoT networks. However, the problem of scalability and adaptability to different types of attacks remains. The use of reinforcement learning-based IDS in adaptive intrusion detection was studied by Palaniappan et al. [11]. The proposed approach helps to detect new and zero-day attacks through learning. But it is not applicable in the real-world scenario because of high computational complexity and complicated reward setting. A new IDS for MANETs based on machine learning and optimization techniques was introduced by Walling et al. [12] to detect routing attacks. It helped to improve the accuracy of detection in a decentralized environment. However, it is not universal and applicable to other network architectures.

Li et al. [13] investigated the effect of adversarial machine learning attacks on IDS models and reported the vulnerability of classifiers like SVM and Random Forest in the presence of such attacks. This paper emphasized the need for a robust defense strategy against such attacks and the necessity of adversarial training. However, it didn't provide any IDS framework. Lee et al. [14] proposed a hybrid approach to feature selection and Random Forest classification-based IDS. This IDS was scalable and efficient in high traffic conditions. However, the utilization of fixed thresholds reduces adaptability and increases false positive cases. Santika et al. [15] proposed a feature selection technique known as ensemble, which was a combination of several techniques used to improve stability and increase accuracy. Overfitting problems were reduced using this technique. However, there was no discussion regarding adaptive thresholding or decision-making.

3. Proposed Work

This research work presents an IDS utilizing the machine learning approach, which involves hybrid feature selection with adaptive threshold tuning technique to achieve better detection performance and reliability. The proposed system will be tested against the UNSW-NB15 dataset [16], which captures realistic network traffic patterns and attacks in order to verify the IDS's performance under realistic settings.

The process of reducing feature dimensions will be done using a hybrid feature selection algorithm, which will make use of the Chi-Squared statistical filter along with Particle Swarm Optimization (PSO). The Chi-Squared will be initially used to remove all the irrelevant features, while the PSO will work to optimize the set of features to enhance the classification performance while decreasing the computational cost. The combination of these two will successfully delete all the irrelevant and redundant features and keep those features/characteristics that are useful in distinguishing the traffic. Supervised Machine Learning algorithms will be used to model both the normal and abnormal traffic behaviours. To increase the reliability of the decision making, there is the Adaptive Threshold Testing Algorithm (ATTA), which varies the classification thresholds based on the feedback from the performance of the decision making in terms of false positive and stability under various conditions.

The integration of both feature selection and classifier optimization in the proposed IDS solves the problems of high dimensionality, imbalance and fixed decision boundary that are inherent to traditional IDSs. In this way, the system becomes more accurate, reliable and efficient and therefore applicable for real-time use in contemporary network environments.

4. Methodology

The ATOM-based IDS design suggests to develop an IDS on machine learning approach to increase accuracy of intrusion detection and decrease the rate of false positives. The architecture is developed using a pipeline of several interconnected processes (data pre-processing, feature selection, classification, and adaptive threshold optimization), which are represented in Figure 1 below. The proposed design solves several challenges associated with traditional IDS architectures.

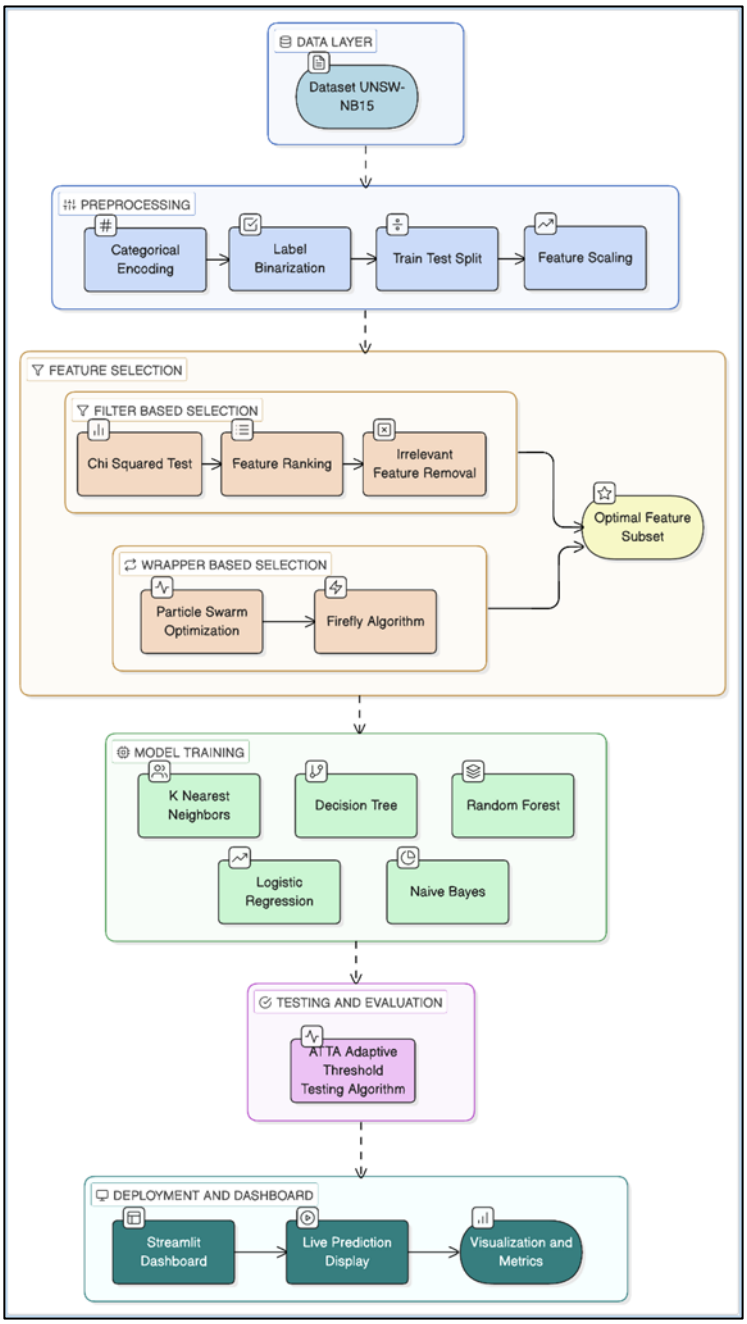

Figure 1. Architecture diagram of the proposed system

4.1 Dataset Collection

The UNSW-NB15 dataset is one of the most popular benchmark datasets that are commonly used in the evaluation of IDSs. This dataset is created by IXIA's PerfectStorm software to generate a simulated network traffic that resembles reality. The UNSW-NB15 dataset provides a variety of both normal and malicious network behaviors, similar to the

operational environment of a network in the real world. In addition, each network instance constitutes a full network behaviour on its own.

The dataset includes:

- Normal and malicious network communications reflecting the realistic network communication patterns.
- Different kinds of cyber-attacks, namely DoS Attacks, Exploits, Reconnaissance, Fuzzing, Shellcode, and Generic.
- 49 features extracted from the network flows consisting of numeric and nominal variables.
- Labeled instances, which are classified either as normal or malicious, thus providing opportunities to apply supervised machine learning algorithms.

Despite the fact that the dataset was created in a synthetic environment, it reflects the network activity and attack patterns in real life. It becomes an efficient tool for evaluating intrusion detection systems' capabilities of detecting and classifying cyber-attacks in the network environment.

4.2 Data Preprocessing

Data preprocessing is an important step to enhance the quality and consistency of the raw network dataset to apply the machine learning algorithms. Preprocessing helps in addressing the issue of missing and inconsistent values in the raw data by either replacing those values with the corresponding statistical measures or simply by eliminating those values to avoid any kind of bias during learning. Due to the presence of categorical attributes like protocol type and service in the data set, label encoding is used to convert these categorical attributes into numeric attributes so as to make them compatible with machine learning algorithms. Normalization is also applied on the numeric attributes to overcome the problem of dominance of attributes with higher magnitude in the learning process.

4.3 Hybrid Feature Selection

The high-dimensional feature space increases the computational overhead and can cause problems in the classification accuracy due to irrelevant and redundant features. In order to solve this problem a combination of feature selection technique is used.

A. Chi-squared Statistical Feature Filtering: Chi-Squared test is employed in order to detect the relationship of each feature to class label. Only the highly dependent features are selected while irrelevant ones are filtered out. This reduces the dimensionality of the data while preserving the discriminating features.

The Chi-Squared score for a feature f is computed as:

$$X^2(f) = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

where O_i is the observed frequency and E_i is the expected frequency:

$$Ei = \frac{(\text{Row Total}) \times (\text{Column Total})}{\text{Grand Total}} \tag{2}$$

B. Particle Swarm Optimization (PSO): In PSO after statistical filtering, feature subset is refined. Each of the particles represents a binary feature vector:

$$X_i = \{X_{i1}, X_{i2}, \ldots.,\ X_{id}\}, X_{ij}\ \in \{0, 1\} \tag{3}$$

Velocity and position are updated iteratively:

$$v_{ij}^{t+1} = wv_{ij}^{t} + c_1\, r_1\big(pbest_{ij} - X_{ij}^{t}\big) + c_2 r_2\big(gbest_j - X_{ij}^{t}\big) \tag{4}$$

$$x_{ij}^{t+1} = \begin{cases} 1, & if\ \sigma\big(v_{ij}^{t+1}\big) > r \\ 0, & otherwise \end{cases} \tag{5}$$

PSO is able to explore efficient subset to find the optimal subsets with the highest detection accuracy while the least number of features.

C. Comparative Evaluation using Firefly Algorithm: The PSO-based optimization scheme is evaluated on the effectiveness through Firefly Algorithm (FA), which is employed as the benchmarking optimization algorithm for the comparison. Feature selections are conducted based on firefly and the obtained feature sets will be compared with those obtained from PSO. However, the ATOM algorithm uses PSO due to its superior performance in detection.

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{6}$$

$$r_{ij} = \sqrt{\sum_{k=1}^{d} \left(x_{ik} - x_{jk}\right)^2} \tag{7}$$

Fireflies converge to better solutions, and enhance the global exploration and prevent early converging.

The proposed hybrid feature selection framework based on Chi-Squared filtering and PSO produces an optimized feature subset that:

- Removes redundant features.
- Improves intrusion detection accuracy.
- Reduces computational complexity.

4.5 Adaptive Threshold Optimization using ATTA

The way to address the shortcomings of upper and lower threshold settings is to apply the algorithm of adaptive threshold testing, ATTA for different cutoff probabilities of sets, which automatically determines the best threshold to be applied with respect to the F1 score optimization and minimization of false positive rates. This type of adaptive technique is known to produce more reliable detection in cases of imbalanced networks.

Working Principle of ATTA

Given network instance x, let the classifier provide a probability value P(x) E [0,1]. Instead of using a fixed threshold, ATTA evaluates a set of candidate thresholds:

$$T = \{t_1, t_2, \dots, t_n\}, t \in [0,1] \tag{8}$$

For each threshold t, the classification decision is defined as:

$$\hat{y}(x) = \begin{cases} 1, & P(x) \geq t \ (Attack) \\ 0, & P(x) < t \ (Normal) \end{cases} \tag{9}$$

Threshold Optimization Criterion

For each threshold t, performance metrics such as false positive rate (FPR) and detection accuracy are computed. The optimal threshold t^* is selected by minimizing false positives while maintaining high accuracy:

$$t^* = arg \min_{t \epsilon T} \big(FPR(t)\big) \, subject \, to \, Accuracy \, (t) \geq \delta \tag{10}$$

ere δ is a minimum acceptable accuracy level.

Alternatively, ATTA can be formulated using a cost-based objective function:

$$Cost(t) = \lambda \cdot FPR(t) + (1 - \lambda) \cdot \big(1 - Accuracy(t)\big) \tag{11}$$

$$t^* = \arg \min_{t} Cost(t) \tag{12}$$

This adaptive strategy allows the IDS to respond dynamically to traffic variations, reduce false alarms, and improve detection reliability, particularly in imbalanced datasets.

Algorithm: Adaptive Threshold Testing Algorithm (ATTA)

Input: Predicted probabilities P(x), threshold set T

Output: Optimal threshold t*

1. Obtain prediction probabilities from trained Random Forest model.
2. Define candidate threshold values:

   T= {0.10, 0.15, …,0.90}
3. For each threshold t ∈ T:
   - Convert probabilities into class labels.
   - Compute Accuracy, Precision, Recall, F1-score, and FPR.
4. Compare all threshold performances.
5. Select threshold maximizing F1-score while minimizing FPR.
6. Assign selected threshold as t*.
7. Use t* for final intrusion detection decisions.
8. Output optimized classification results.

5. Result and Analysis

The proposed ATOM framework is coded in the Python Programming Language and uses the Scikit-learn library to create models and NumPy to handle data and simplify information presentation in the interaction with the user. The modular pipeline is capable of integrating the pre-processing technique, hybrid feature selection, supervised classification,

hybrid threshold adjustment, and monitoring by the dashboard. The experimental model provides the opportunity to compare and optimize classifiers.

A. Chi-squared Feature Selection: Figure 2 illustrates that the statistical method, Chi-Squared, was applied in order to distinguish the attributes into those which are relevant and irrelevant to the attack label. The rank scores were used to select the best 50 features which will significantly reduce the dimensions but yet provide discriminative information.

```
================ CHI2 FEATURE SELECTION =================

Chi2 selected top 50 features
Chi2 feature scores (top 10): [39148.55794412 37268.44769508 36494.52833842 17736.90738523
 17253.41016239 14574.6678534  14573.14142764 13853.03858995
 12260.79517192 10683.82590183]
```

Figure 2. Chi-squared feature selection output

B. Particle Swarm Optimization (PSO): PSO was applied to obtain even better results in feature selection. Particles here are the binary mask of the features and the fitness value used here is the Rand Forest F1-score. Using PSO, the best set of 26 features was found in just 10 iterations, which are shown in figure 3.

```
================ PSO FEATURE SELECTION =========

PSO Iteration 1/10 completed
PSO Iteration 2/10 completed
PSO Iteration 3/10 completed
PSO Iteration 4/10 completed
PSO Iteration 5/10 completed
PSO Iteration 6/10 completed
PSO Iteration 7/10 completed
PSO Iteration 8/10 completed
PSO Iteration 9/10 completed
PSO Iteration 10/10 completed

PSO completed
PSO selected features: 26 out of 50 (Chi2 base)
```

Figure 3. PSO feature optimization output

C. Firefly-based Comparative Analysis: Firefly Algorithm was tested as another algorithm that can be used as an alternative to PSO. Though Firefly algorithm was able to reduce the number of features to 11 features, PSO provided better classification accuracy and detection ability. Thus, the final ATOM algorithm employs the 26 features chosen by PSO.

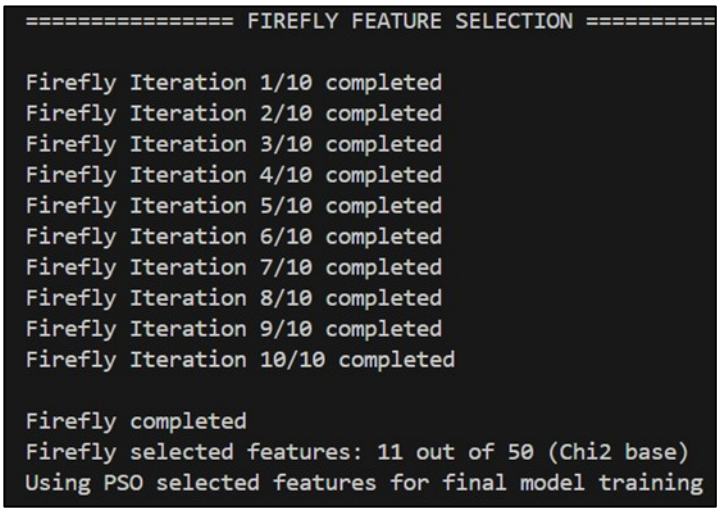

Figure 4. Firefly feature selection output

5.1 Adaptive Threshold Testing Algorithm (ATTA)

To improve the accuracy of detection, the Adaptive Threshold Testing algorithm is used, thus making use of the dynamic threshold for setting up an optimal probability threshold instead of a static probability threshold. Thresholds are evaluated and the one that gives the highest F1-score while maintaining the lowest false positive is chosen. The ideal threshold is 0.31, leading to an F1-score of 0.9609 and reducing false alarms is shown in Table 2.

Table 2. ATTA optimization and final performance results

Parameter	Value
Optimal Threshold (ATTA)	0.31
Best F1-Score (After ATTA)	0.9609
Accuracy	0.9756
Precision	0.9167
Recall	0.9682
False Positive Rate (FPR)	0.0137

5.2 Streamlit Dashboard

The trained IDS models are applied using the interactive Streamlit dashboard to continue observing the visualization and performance. The interface offers capability to do the following: select the model, change threshold, observe key performance indicators: accuracy,

precision, recall, F1-score, and false positive rate. Additional options include visualization of confusion matrix, ROC plot, probability of an attack, traffic distributions graph and alerts on high risk network traffic.

A. Classification Analysis Dashboard: Figure 5 shows the performance of the intrusion detection system based on the confusion matrix, flow classification distribution, and ROC curve. These graphs help assess the effectiveness of the model for distinguishing between normal and malicious network traffic.

Figure 5. Classification analysis showing confusion matrix, traffic distribution, and ROC curve.

B. Model Performance Comparison: This section shows key evaluation metrics such as accuracy, precision, recall, F1-score, false positive rate and AUC (Figure 6). Radar chart and comparison table are used to compare the performance of various machine learning models such as Random Forest, Logistic Regression, Decision Tree, KNN, and Naïve Bayes.

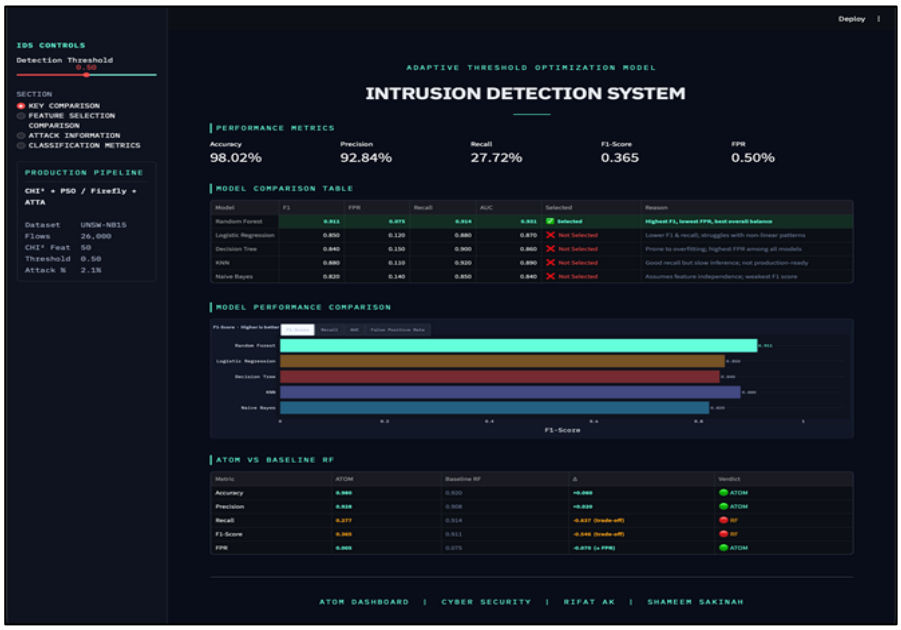

Figure 6. Machine learning model performance comparison using radar visualization and metric table.

C. Feature Selection Comparison: The feature selection dashboard (Figure 7) is used for comparing various optimization algorithms such as the Chi-Squared algorithm, Particle Swarm Optimization (PSO), and Firefly Algorithm. The dashboard shows the important features selected using the optimized IDS pipeline.

Figure 7. Feature selection comparison and visualization of top important network traffic features.

D. Attack Information and Alerts: Figure 8 displays the attack information and alerts obtained from the analysis of network traffic data. It contains information regarding the types of attacks, most critical alerts, and probability of attacks through attack probability wave.

Figure 8. Attack detection dashboard showing attack distribution, alert table, and probability analysis

6. Conclusion and Future Scope

This study has discussed the Adaptive Threshold Optimization Model (ATOM). It is a machine learning-based IDS that is able to improve detection accuracy and reduce the number of false positives at the same time. The use of a hybrid feature selection technique involving Chi-Squared filtering and Particle Swarm Optimization (PSO) resulted in reducing the feature set from 138 to only 26 significant features without compromising on the results obtained. Random Forest classifier performed better than other classifiers because of its stability, ensemble technique, and ability to deal with non-linear relationships. Additionally, the use of the Adaptive Threshold Testing Algorithm (ATTA) helped to find the optimal value of the threshold for classification and achieved a better precision-recall ratio. Based on the experimental study with the UNSW-NB15 data set, we obtained an F1-score equal to 0.4409 and the low false-positive rate equal to 1.37%.

References

[1] Emirmahmutoğlu, Emre, and Yılmaz Atay. "A feature selection-driven machine learning framework for anomaly-based intrusion detection systems." Peer-to-Peer Networking and Applications 2025, vol. 18, no. 3: 161.

[2] Kumar, Neeraj, and Sanjeev Sharma. "A hybrid modified deep learning architecture for intrusion detection system with optimal feature selection." Electronics 2023, vol. 12, no. 19: 4050.

[3] Bakır, Halit, and Özlem Ceviz. "Empirical enhancement of intrusion detection systems: a comprehensive approach with genetic algorithm-based hyperparameter tuning and hybrid feature selection." Arabian Journal for Science and Engineering 2024, vol.49, no. 9: 13025-13043.

[4] Pattawaro, Apichit, and Chantri Polprasert. "Anomaly-based network intrusion detection system through feature selection and hybrid machine learning technique." 16th International Conference on ICT and Knowledge Engineering (ICT&KE) 2018, IEEE, 1-6.

[5] Turukmane, Anil V., and Ramkumar Devendiran. "M-MultiSVM: An efficient feature selection assisted network intrusion detection system using machine learning." Computers & Security 2024, vol.137: 103587.

[6] Kareem, Saif S., Reham R. Mostafa, Fatma A. Hashim, and Hazem M. El-Bakry. "An effective feature selection model using hybrid metaheuristic algorithms for iot intrusion detection." Sensors 2022, vol. 22, no. 4: 1396.

[7] Alkanhel, Reem, El-Sayed M. El-kenawy, Abdelaziz A. Abdelhamid, Abdelhameed Ibrahim, Manal Abdullah Alohali, Mostafa Abotaleb, and Doaa Sami Khafaga. "Network Intrusion Detection Based on Feature Selection and Hybrid Metaheuristic Optimization." Computers, Materials & Continua 2023, vol, 74, no. 2.

[8] Ahmed, Naveed, Bilal Zahran, Belal Ayyoub, Abdel Rahman Alzoubaidi, and Md Asri Ngadi. "Intrusion Detection System Using Hybrid Machine Learning Classifiers and

Optimum Feature Selection in Internet of Things (IoT)." International Journal on Communications Antenna & Propagation 2024, vol. 14, no. 2: 51.

[9] Megantara, Achmad Akbar, and Tohari Ahmad. "A hybrid machine learning method for increasing the performance of network intrusion detection systems." Journal of Big Data 2021, vol. 8, no. 1: 142.

[10] Velliangiri, S., and P. Karthikeyan. "Hybrid optimization scheme for intrusion detection using considerable feature selection." Neural Computing and Applications 2020, vol. 32, no. 12: 7925-7939.

[11] Palaniappan, Saranya, and Sudhakar Sengan. "Hybrid feature selection for IoMT based intrusion detection system for integrating mutual information filtering with deep learning based accelerated metaheuristic optimization." Scientific Reports (2026).

[12] Walling, Supongmen, and Sibesh Lodh. "Network intrusion detection system for IoT security using machine learning and statistical based hybrid feature selection." Security and Privacy 2024, vol. 7, no. 6: e429.

[13] Li, Jing, Mohd Shahizan Othman, Hewan Chen, and Lizawati Mi Yusuf. "Optimizing IoT intrusion detection system: feature selection versus feature extraction in machine learning." Journal of Big Data 2024, vol. 11, no. 1: 36.

[14] Lee, Yu-Rim, So-Eun Jeon, Sun-Jin Lee, and Il-Gu Lee. "Adaptive feature selection with gradient-based relevance for intrusion detection systems." Scientific Reports (2026).

[15] Santika, Satriadi Putra, Ford Lumban Gaol, Yulyani Arifin, and Haryono Soeparno. "Evaluating Hybrid Feature Selection Quality and Efficiency Trade-Off in DRL-Based Intrusion Detection Systems." 8th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, 2025, 1066-1071.

[16] Intrusion Detection Dataset - https://www.kaggle.com/datasets/mrwellsdavid/unsw-nb15