Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(3) June 2026 Special Issue
DOI: https://dx.doi.org/10.4314/jobasr.v4i3.6s

Hybrid Ant Colony-Cuttlefish Optimization for Feature Selection in Machine Learning-Based Intrusion Detection Systems

Opeyemi Lateef USMAN[1*]
[1]Department of Computer Science, Tai Solarin Federal University of Education, Ijagun, Ogun State, Nigeria.
[*]Corresponding Author Email: usmanol@tasued.edu.ng

ABSTRACT
Intrusion Detection Systems (IDS) constitute a critical component of contemporary cybersecurity frameworks; however, their performance is frequently hindered by the high dimensionality of network traffic data, which contributes to increased false alarm rates and computational inefficiencies in conventional feature selection approaches. To address these limitations, this study proposes a novel hybrid feature selection framework that integrates the Ant Colony Optimization (ACO) and the Cuttlefish Algorithm (CFA). The hybrid approach leverages the global search exploration capability of ACO alongside the local search refinement strength of CFA to improve feature optimization for intrusion detection tasks. The proposed ACO-CFA algorithm was extensively evaluated using the KDD Cup 99 benchmark dataset, while the effectiveness of the optimized feature subsets was assessed through three machine learning classifiers: Random Forest (RF), Decision Tree (DT), and Support Vector Machine (SVM). Experimental findings revealed outstanding classification performance across multiple data partitioning ratios. Among the evaluated models, the Random Forest classifier consistently demonstrated superior effectiveness, achieving an accuracy of up to 99.97% while maintaining an optimal balance between precision and recall. Although the DT classifier produced comparable accuracy with faster computational performance, the SVM exhibited substantially higher computational costs despite its strong predictive capability. The study concludes that the hybrid ACO-CFA framework provides an efficient and scalable solution for feature selection, significantly improving IDS detection accuracy while minimizing computational complexity. Consequently, the proposed approach offers a robust foundation for the development of adaptive and high-performance intrusion detection systems capable of addressing increasingly sophisticated cyber threats.

Keywords:
Intrusion Detection System,
Hybrid Feature Selection,
Ant Colony Optimization,
Cuttlefish Algorithm,
Cybersecurity

INTRODUCTION

Intrusion Detection Systems (IDS) constitute a fundamental component of modern cybersecurity frameworks, serving as essential mechanisms for safeguarding networks, systems, and sensitive data against increasingly sophisticated cyber threats. These systems continuously monitor network traffic, user activities, and system behaviors with the objective of identifying unauthorized access attempts, malware intrusions, anomalous activities, and other forms of security breaches. By providing timely detection and alert mechanisms, IDS contribute significantly to maintaining the confidentiality, integrity, and availability of information systems in both organizational and enterprise environments (Kannan, 2020; Prithi & Sumathi 2024; Panliang et al. 2025).

Despite their critical role in cybersecurity, the effectiveness of conventional IDS has been increasingly undermined by the rapidly evolving and sophisticated nature of modern cyberattacks (Akoul et al., 2026; Kamil et al., 2022; Radhakrishnan et al., 2019). Contemporary cyber threats are more adaptive, intelligent, and difficult to detect, thereby exposing significant limitations in traditional IDS feature selection methodologies. In particular, many existing approaches are characterized by reduced detection accuracy, elevated false alarm rates, high computational complexity, and inefficient processing of large-scale network traffic data (Alwan et al., 2020; Ogundokun et al., 2022; Panliang et al., 2025). These challenges negatively impact the reliability, scalability, and real-time responsiveness of IDS frameworks in practical cybersecurity environments.

How to cite this article: Opeyemi L. U. (2026). Hybrid Ant Colony-Cuttlefish Optimization for Feature Selection in Machine Learning-Based Intrusion Detection Systems. *Journal of Basics and Applied Sciences Research*, *4*(3), 48-58. https://dx.doi.org/10.4314/jobasr.v4i3s.6

Moreover, modern IDS datasets typically exhibit high dimensionality, often comprising numerous redundant, irrelevant, or noisy features that degrade model performance. The presence of such extraneous attributes increases computational burden, extends processing time, and reduces classification efficiency, ultimately weakening the overall predictive capability of intrusion detection models (Kamil et al., 2022; Rufai et al., 2016; Sandhya et al., 2024). Consequently, there is a pressing need for more advanced and intelligent feature selection techniques capable of isolating the most informative attributes while effectively eliminating redundancy. In this context, the integration of metaheuristic optimization algorithms with machine learning and hybrid intelligent frameworks has emerged as a promising research direction for enhancing IDS performance (Balasaraswathi & Sugumaran, 2019; Safana et al. 2026). Such approaches have been shown to improve detection accuracy, reduce false positive rates, and strengthen the efficiency and robustness of cyber threat detection in complex and dynamic network environments (Rufai et al., 2021; Saheed, 2022).

This study proposes a hybrid feature selection approach that integrates Ant Colony Optimization (ACO) and the Cuttlefish Algorithm (CFA) to address the limitations associated with traditional IDS. Ant Colony Optimization, inspired by the pheromone-guided foraging behavior of ants, is highly effective in exploring large and complex search spaces to identify the most informative and relevant features within high-dimensional datasets. In contrast, the Cuttlefish Algorithm emulates the adaptive color-changing mechanisms of cuttlefish, enabling the refinement and optimization of selected feature subsets to achieve improved classification performance (Aghdam & Kabiri, 2016; Eesa et al. 2015; Eesa et al. 2021). The integration of these two bio-inspired optimization techniques is intended to leverage the exploration strength of ACO alongside the exploitation and refinement capabilities of CFA. Through this hybridization, the proposed model seeks to eliminate redundant and irrelevant features, thereby reducing data dimensionality and computational complexity. Furthermore, the approach is designed to enhance the detection accuracy of the machine learning-based IDS, minimize false alarm rates, and improve the overall efficiency of cyberattack detection and classification processes (Akoul et al. 2026; Balasaraswathi et al., 2018; Balasaraswathi et al., 2017). Consequently, the hybrid ACO-CFA framework provides a robust and intelligent solution for strengthening IDS performance in modern cybersecurity environments characterized by increasingly sophisticated and high-volume cyber threats.

MATERIALS AND METHODS

Data Collection and Preprocessing

Data collection plays a crucial role in the development of an effective intrusion detection system (IDS), as the quality and relevance of the dataset significantly influence model performance. In this study, the KDD Cup 99 dataset was utilized due to its extensive feature set and comprehensive representation of real-world cyberattack scenarios. The dataset contains 42 attributes and 494,017 records encompassing both normal and malicious network traffic, including Denial-of-Service (DoS), Probe, User-to-Root (U2R), and Remote-to-Local (R2L) attacks.

To enhance the reliability and efficiency of the intrusion detection model, a comprehensive preprocessing phase was conducted. This involved data cleaning techniques such as handling missing values, removing duplicate records, and eliminating features with excessive missing data. Feature transformation methods, including normalization and standardization defined in Eq. (1)-(2), were also applied to ensure uniform scaling of numerical attributes and improve the learning capability of the machine learning model. These preprocessing procedures contributed to improving data quality, reducing noise, and enhancing the overall performance of the proposed intrusion detection framework.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

$$X' = \frac{X - \mu}{\sigma} \quad (2)$$

where X is the original value, X' is the transformed value, μ is the mean, and $\boldsymbol{\sigma}$ is the standard deviation.

Proposed System Architecture

The proposed system architecture, as presented in Figure 1, was designed to enhance intrusion detection efficiency through a hybrid feature selection approach that integrates Ant Colony Optimization (ACO) and the Cuttlefish Algorithm (CFA). The framework consists of interconnected stages, including data collection and preprocessing, feature selection, feature optimization, intrusion classification, and performance evaluation. Initially, network traffic data obtained from source, specifically, the KDDCUP99 dataset undergo preprocessing operations including cleaning, normalization, transformation, and feature encoding to improve data quality and suitability for machine learning models.

Subsequently, the ACO algorithm is employed to identify the most relevant features by simulating the pheromone-guided foraging behavior of ants (Rufai et al. 2021), while the CFA further refines the selected feature subsets through an optimization process inspired by cuttlefish color adaptation mechanisms. The optimized feature set is then utilized in the classification stage, where machine learning classifiers such as Support Vector Machine (SVM), Decision Tree, and Random Forest are trained to distinguish between normal and malicious network activities. Finally, the effectiveness of the IDS is

evaluated using performance metrics including accuracy, precision, recall, and F1-score to determine its reliability and suitability for real-time cybersecurity applications.

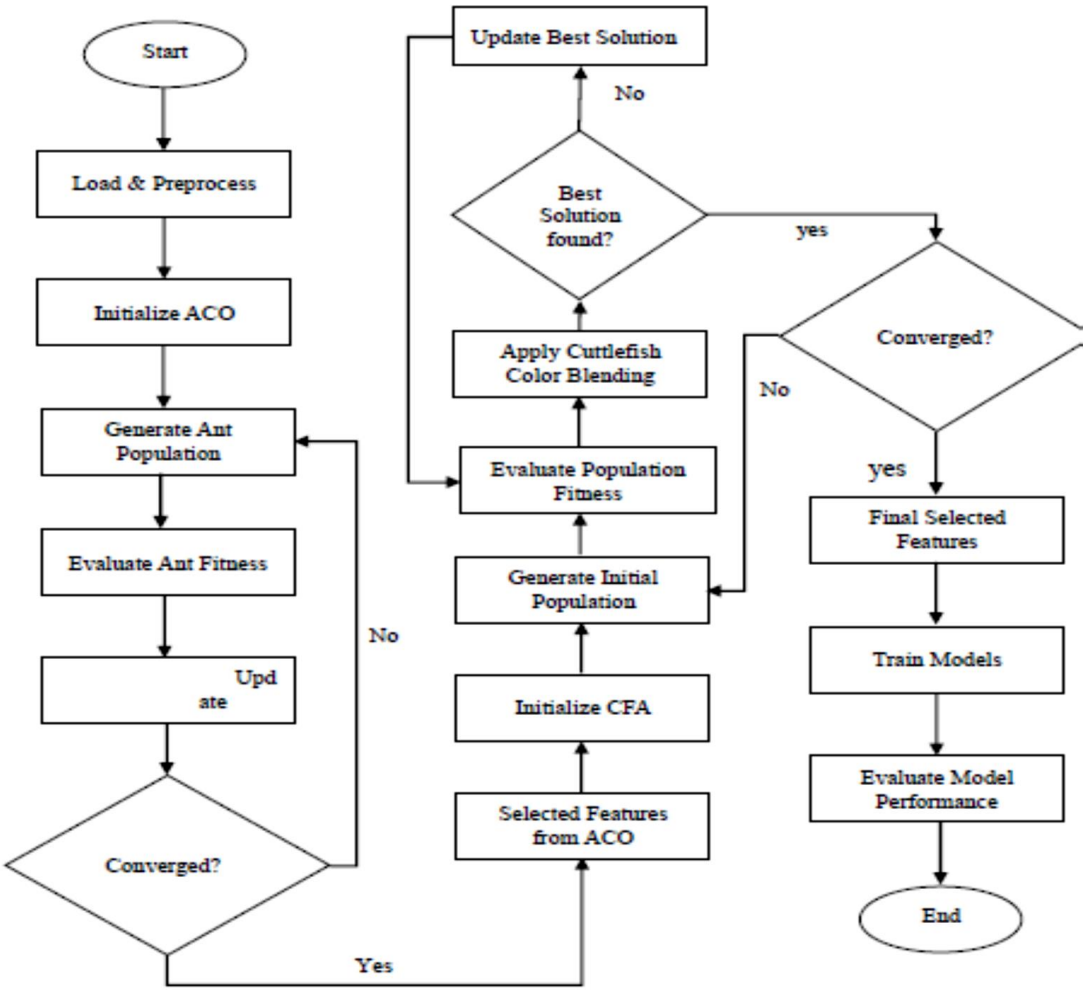


Figure 1. Proposed system architecture flowchart integrating ACO-CFA framework.

Machine Learning Models for Intrusion Detection and Classification

The refined feature subsets were utilized to train and evaluate three machine learning classifiers, namely Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF). These algorithms were selected due to their strong generalization capabilities and their efficiency in processing high-dimensional datasets. Through the learning process, the models were able to effectively distinguish between normal and malicious network activities, thereby achieving accurate intrusion classification.

Among the selected classifiers, SVM is a supervised learning technique that classifies data by constructing an optimal hyperplane within an *N*-dimensional feature space. The preprocessed training dataset was provided to the SVM model using an appropriate kernel function to enhance classification performance. The model determines the optimal separating hyperplane by maximizing the margin between different classes. Furthermore, hyperparameter optimization was carried out using Grid Search Cross-Validation to identify the most suitable values for the regularization parameter (*C*) and kernel coefficient (gamma), based on the decision function expressed in Eq. (3) (Ogundokun et al. 2022; Usman et al. 2021).

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \quad (3)$$

where α_i are the Lagrange multipliers, y_i are the class labels, $K(x_i, x)$ is the kernel function, and b is the bias term.

The Decision Tree (DT) is a rule-based supervised machine learning algorithm that classifies data by

partitioning instances according to feature values, thereby generating a hierarchical tree-like structure. The algorithm recursively divides the dataset into smaller subsets to establish a sequence of decision rules, making the model highly interpretable and easy to understand (Usman et al. 2021). The training process of the DT model involves the implementation of the ID3 (Iterative Dichotomiser 3) algorithm, which recursively performs data splitting until specified stopping criteria are satisfied. To minimize overfitting and improve model generalization, pruning strategies such as pre-pruning, through maximum depth restriction, and post-pruning, through the elimination of redundant nodes, were employed in accordance with the decision function presented in Eq. (4) (Yinka-Banjo et al. 2022).

$$IG(T,X) = H(T) - \sum_{v \in V} P(v)H(T_v) \quad (4)$$

where *IG(T,X)* is the information gain of feature *X*, *H(T)* is the entropy of the dataset *T*, *P(v)* is the probability of each value v, *H(T_v)* is the entropy of the subset T_v after splitting by *X*.

The Random Forest (RF) classifier is an ensemble-based machine learning technique that operates by constructing multiple decision trees and combining their outputs to produce a final prediction. This ensemble strategy improves classification accuracy while simultaneously reducing the risk of overfitting. The training process of the RF model involves the parallel development of several DT models using different subsets of the training dataset. Each individual tree generates an independent prediction, while the final classification outcome is determined through a majority voting mechanism (Usman et al. 2021; Usman et al. 2025a). To enhance model performance, hyperparameter optimization was conducted on key parameters, including the number of trees (*n*_estimators), which determines the size of the ensemble; the maximum tree depth, which restricts excessive tree growth to mitigate overfitting; and the minimum number of samples required for splitting, which helps prevent excessive partitioning of smaller datasets. The decision function governing the RF model is expressed in Eq. (5) (Zivkovic et al. 2022).

$$F(x) = \frac{1}{N}\sum_{i=1}^{N} T_i(x) \quad (5)$$

where *N* is the number of trees in the forest, *T_i(x)* is the prediction of the *i*-th tree for input *x*, *F(x)* is the final prediction based on majority voting.

Experimental Setup and Implementation Configuration

All experimental procedures were carried out using the Python programming language, supported by the TensorFlow and Keras libraries. In particular, Scikit-learn was employed for the modeling, training, and evaluation of the machine learning algorithms, while NumPy and Pandas facilitated data preprocessing and manipulation. Additionally, Matplotlib was utilized for the visualization of experimental results. To improve computational efficiency and accelerate processing speed, model training and optimization were executed within a GPU-enabled computing environment (Usman et al., 2025; Usman & Adeusi 2025; Usman et al. 2026). The adopted technical setup aligns with standard experimental frameworks commonly used in intrusion detection system research and reflects the computational and resource limitations highlighted in energy-efficient modeling studies, including those reported by Usman and Muniyandi (2020) and Usman et al. (2021). The parameter settings applied during the simulation and implementation of the hybrid Ant Colony Optimization and Cuttlefish Algorithm (ACO-CFA) framework for optimal feature selection are summarized below:

1. Number of ants in ACO = 100
2. Number of ACO iterations = 200
3. Pheromone evaporation rate = 0.6
4. Population size for CFA = 50
5. Number of CFA iterations = 200
6. Mutation rate for CFA = 0.1

Performance Evaluation Metrics

The confusion matrix was employed to assess the classification performance of the trained models, including the Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) classifiers. It offers a detailed representation of correct and incorrect predictions by comparing the actual class labels with the predicted outcomes. Based on the confusion matrix, essential evaluation metrics such as accuracy, precision, recall, and F1-score were computed. These performance indicators provide comprehensive insights into the effectiveness of the models in accurately detecting intrusions while reducing the occurrence of false positives and false negatives. The metrics are described as follows:

1. **Accuracy:** Accuracy is a fundamental evaluation metric used to measure the overall performance of a classification model by determining the proportion of correctly classified instances relative to the total dataset. Although a high accuracy value generally indicates strong model performance, it may not provide sufficient evaluation in cases of imbalanced datasets, thereby necessitating the use of additional performance metrics. This metric is defined using Eq. (6).
$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$
2. **Precision:** Precision evaluates the proportion of correctly identified attack instances among all predicted attack cases, reflecting the accuracy of positive predictions. A high precision value indicates a low false positive rate, which is essential in intrusion detection systems to

minimize unnecessary alerts on legitimate network activities. This metric is defined using Eq. (7).

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

3. **Recall (Detection Rate):** Recall, also referred to as Detection Rate (DR) or Sensitivity, measures the proportion of actual attack instances correctly identified by the model. A high recall value indicates the model's effectiveness in detecting intrusions and minimizing the likelihood of undetected attacks. This metric is defined using Eq. (8).

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

4. **F1-Score:** The F1-score represents the harmonic mean of precision and recall, offering a balanced measure of both metrics. It is particularly valuable in the evaluation of imbalanced datasets, where class distributions are unequal. A higher F1-score reflects a model that effectively balances false positives and false negatives, indicating robust overall performance. This metric is defined using Eq. (9).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

5. **Computational Time:** Beyond classification performance, computational efficiency is a critical requirement in real-time intrusion detection systems (IDS). Computational time refers to the total duration required for feature selection, model training, and classification processes. Minimizing computational time is essential to ensure efficient processing of network traffic and to support timely detection and response, thereby enhancing the feasibility of deployment in real-world cybersecurity environments.

Convergence Curve: This metric provides a visual representation of how an algorithm's performance evolves and stabilizes over successive iterations or epochs. It is useful for assessing convergence behaviour by indicating the rate at which the algorithm attains an optimal or satisfactory solution, as well as its stability, reflected in consistent performance without significant fluctuations.

RESULTS AND DISCUSSION

The performance evaluations of the proposed Intrusion Detection System (IDS), which integrates the hybrid Ant Colony Optimization (ACO) and Cuttlefish Algorithm (CFA) framework for optimized feature selection with machine learning-based classification, are presented in Tables 1–5. To comprehensively assess the robustness, stability, and generalization capability of the proposed framework under different dataset configurations, a series of experiments were conducted using five distinct training-to-validation/testing data partition ratios: 85:15, 80:20, 75:25, 70:30, and 65:35. These partitioning strategies were adopted to investigate the impact of varying training data sizes on the predictive performance and reliability of the IDS models. The experimental results consistently demonstrated that the hybrid ACO-CFA framework effectively enhanced feature selection by identifying the most informative and discriminative attributes while eliminating redundant and irrelevant features. Consequently, the machine learning classifiers trained on the optimized feature subsets achieved remarkably high detection accuracy, reduced false positive and false negative rates, and exhibited strong adaptability across different network intrusion scenarios. Furthermore, the framework maintained stable performance across all evaluated partition ratios, indicating its robustness and effectiveness in handling variations in training and testing data distributions.

A comparative analysis of the results revealed that the 80:20 training-to-validation/testing partition ratio produced the most balanced and optimal performance across the evaluation metrics, including accuracy, precision, recall, F1-score, ROC-AUC value, and computational time. This partition ratio provided an effective trade-off between model learning and generalization, enabling the classifiers to achieve superior detection capability while minimizing overfitting. As a result, the performance obtained using the 80:20 split was selected as the benchmark result for this study and served as the primary basis for evaluating and discussing the effectiveness of the proposed hybrid ACO-CFA-based IDS framework. The findings further confirm the suitability of the proposed approach for developing accurate, efficient, and reliable intrusion detection systems capable of addressing contemporary cybersecurity challenges.

Table 1. Performance of IDS-based machine learning classifiers using partition 85:15.

Classifier	Accuracy (%)	Class	Precision (%)	Recall (%)	F1-Score (%)	Computational Time (mins)
SVM	99.58	Normal	98.06	99.89	98.97	31.6
		Attack	99.97	99.52	99.74	
DT	99.92	Normal	99.87	99.88	99.88	0.39
		Attack	99.97	99.97	99.97	

RF	99.95	Normal	99.90	99.97	99.93	4.7
		Attack	99.39	99.87	99.94	

Table 2. Performance of IDS-based machine learning classifiers using partition 80:20.

Classifier	Accuracy (%)	Class	Precision (%)	Recall (%)	F1-Score (%)	Computational Time (mins)
SVM	99.55	Normal	97.92	99.89	98.97	21.2
		Attack	99.95	99.98	99.79	
DT	99.96	Normal	99.92	99.89	99.90	0.32
		Attack	99.98	99.98	99.97	
RF	99.97	Normal	99.95	99.95	99.94	4.4
		Attack	99.99	99.97	99.98	

Table 3. Performance of IDS-based machine learning classifiers using partition 75:25.

Classifier	Accuracy (%)	Class	Precision (%)	Recall (%)	F1-Score (%)	Computational Time (mins)
SVM	99.47	Normal	97.51	99.82	98.68	8.6
		Attack	99.91	99.37	99.67	
DT	99.95	Normal	99.87	99.91	99.89	0.31
		Attack	99.97	99.96	99.95	
RF	99.97	Normal	99.90	99.95	99.93	4.3
		Attack	99.98	99.95	99.98	

Table 4. Performance of IDS-based machine learning classifiers using partition 70:30.

Classifier	Accuracy (%)	Class	Precision (%)	Recall (%)	F1-Score (%)	Computational Time (mins)
SVM	99.79	Normal	99.08	99.85	99.46	109.2
		Attack	99.92	99.77	99.86	
DT	99.95	Normal	99.90	99.87	99.88	0.29
		Attack	99.93	99.92	99.95	
RF	99.95	Normal	99.92	99.96	99.94	4.4

Table 5. Performance of IDS-based machine learning classifiers using partition 65:35.

Classifier	Accuracy (%)	Class	Precision (%)	Recall (%)	F1-Score (%)	Computational Time (mins)
SVM	99.55	Normal	97.85	99.92	98.87	59.3
		Attack	99.98	99.46	99.72	
DT	99.95	Normal	99.90	99.88	99.89	0.42
		Attack	99.97	99.97	99.97	
RF	99.97	Normal	99.90	99.96	99.93	4.0
		Attack	99.99	99.97	99.98	

According to Table 2, the Support Vector Machine (SVM) model correctly classified 79,029 attack instances and 19,334 normal instances, recording 409 false negatives and 32 false positives. Similarly, the Decision Tree (DT) classifier accurately identified 79,424 attack instances and 19,342 normal instances, with only 14 false negatives and 24 false positives. The Random Forest (RF) classifier demonstrated superior performance by correctly

classifying 79,418 attack instances and 19,357 normal instances, while maintaining a minimal error rate of 20 false negatives and 9 false positives. Furthermore, the Area under the Curve (AUC) values for the three models ranged between 0.96 and 0.99, indicating excellent discriminatory capability and class separability. Based on these findings, the RF classifier emerged as the best-performing model, achieving superior results across key evaluation metrics, including accuracy, precision, recall, and F1-score. Figure 2 presents the confusion matrices and ROC–AUC curves for the three selected machine learning models.

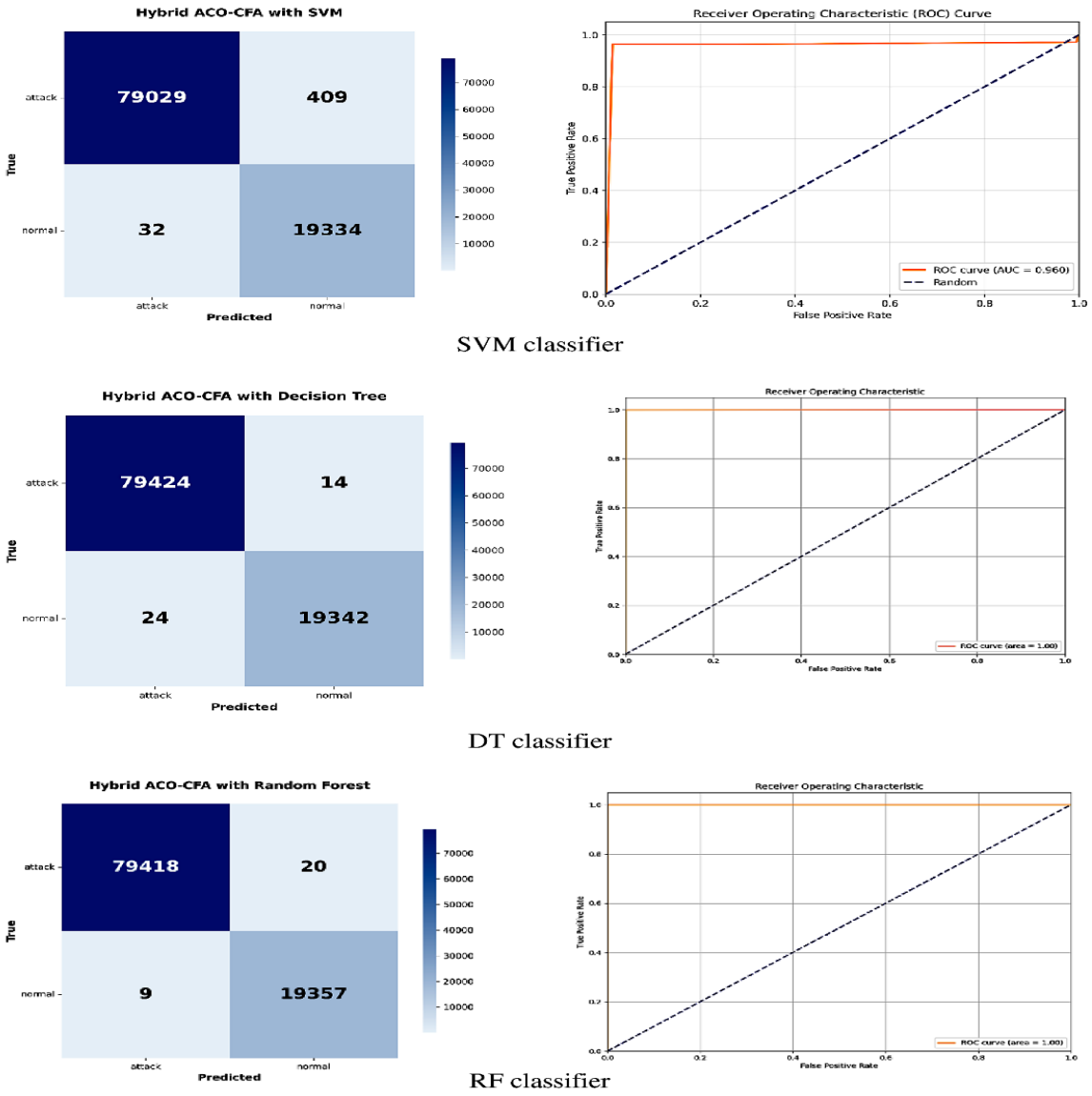
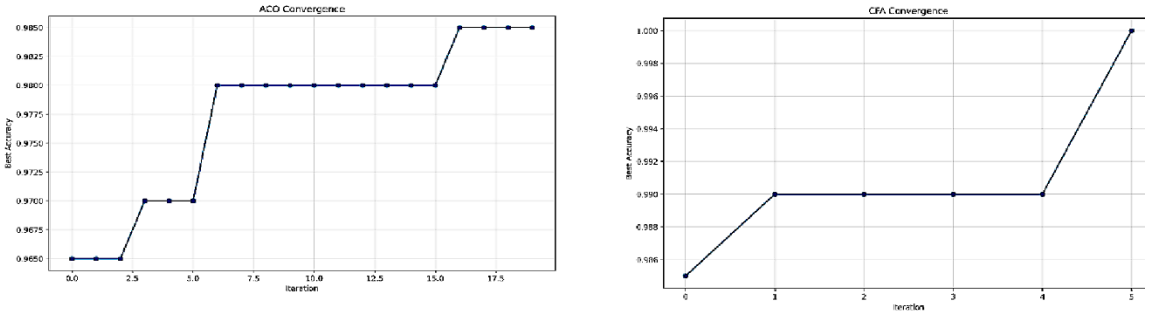
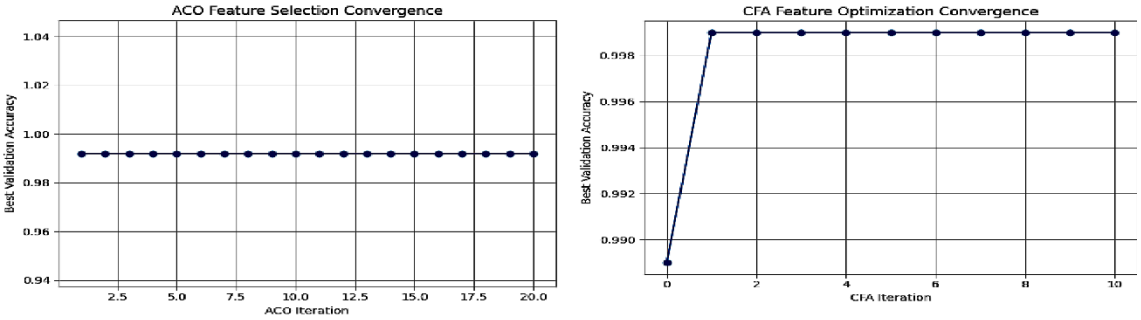


Figure 2. Confusion matrices and ROC-AUC curves of machine learning models.

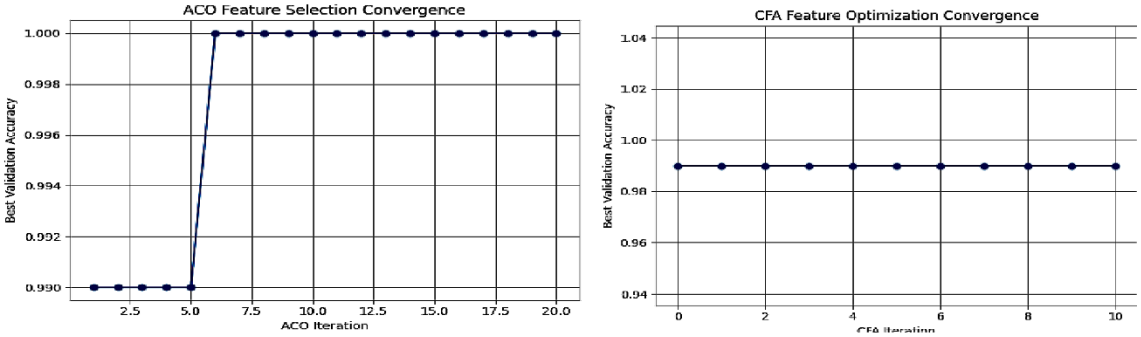
Figure 3 presents the convergence curves of the hybrid Ant Colony Optimization–Cuttlefish Algorithm (ACO-CFA) framework for optimized feature selection across the three machine learning classifiers used in the intrusion detection system (IDS). The results indicate that the hybrid framework achieved rapid convergence and high classification accuracy with strong stability across iterations.



Convergence curve of hybrid ACO-CFA framework for SVM classifier



Convergence curve of hybrid ACO-CFA framework for DT classifier



Convergence curve of hybrid ACO-CFA framework for RF classifier

Figure 3. Convergence curves of the hybrid ACO-CFA framework.

For the Support Vector Machine (SVM), the ACO process demonstrated gradual improvements in accuracy before converging at 98.50%, while the CFA stage maintained stable performance, suggesting that the selected feature subset was already near-optimal. In the Decision Tree (DT) model, ACO maintained consistently high accuracy throughout the iterations, whereas CFA quickly improved and stabilized performance at approximately 99.85%. For the Random Forest (RF) classifier, both ACO and CFA achieved and sustained perfect accuracy (100%) from the initial iterations, indicating immediate convergence to an optimal feature subset. Overall, the convergence behavior demonstrates the effectiveness of the hybrid ACO-CFA framework in identifying highly relevant features, thereby contributing to robust and accurate intrusion detection performance.

This section provides a comprehensive evaluation of the proposed intrusion detection system (IDS), which integrates a hybrid Ant Colony Optimization–Cuttlefish Algorithm (ACO-CFA) framework for feature optimization with Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) classifiers. Performance assessment was conducted using multiple training-to-testing partition ratios and evaluation metrics, including accuracy, precision, recall, F1-score, ROC–AUC, computational time, and convergence behavior. The findings demonstrate the strong effectiveness of the

proposed system, highlighting both the efficiency of the hybrid feature selection mechanism and the high predictive capability of the classification models.

The convergence analysis revealed that the hybrid ACO-CFA framework effectively identified optimal feature subsets through a two-stage optimization process, where ACO rapidly explored the feature space and CFA refined the selected features to achieve stable and near-optimal performance. This optimized feature selection contributed significantly to the consistently high classification outcomes observed across all models. The classifiers achieved exceptional performance, with accuracy values generally exceeding 99.5%, while ROC–AUC results indicated near-perfect or perfect class separability, particularly for DT and RF classifiers. Among the evaluated models, the RF classifier consistently demonstrated superior overall performance, achieving the highest accuracy, precision, recall, and F1-score, alongside strong robustness and acceptable computational cost. The DT classifier also produced highly reliable results with the added advantages of computational efficiency and interpretability, whereas the SVM model, despite strong detection capability, exhibited greater computational demands and variability in training time.

Furthermore, varying dataset partition ratios confirmed the stability and generalization ability of the RF and DT classifiers, as their performance remained consistently high across different training and testing configurations. Computational efficiency analysis showed DT as the fastest model, RF as moderately efficient with balanced performance, and SVM as the most computationally intensive despite optimization measures. Overall, the findings validate the effectiveness of the hybrid ACO-CFA framework in enhancing intrusion detection performance and suggest that RF provides the most balanced solution for practical deployment, while DT offers a suitable alternative in scenarios prioritizing speed and interpretability.

Table 6 provides a comparative evaluation of the proposed hybrid framework against a few selected state-of-the-art feature selection approaches previously applied in IDS studies. The comparative results demonstrate that the proposed hybrid Ant Colony Optimization-Cuttlefish Algorithm (ACO-CFA) framework achieves superior intrusion detection performance, achieving an overall detection rate of 99.97%. Specifically, the traditional Ant Colony Optimization (ACO)-based feature selection approach proposed by Aghdam & Kabiri (2016) attained a detection rate of 98.90%. Likewise, Balasaraswathi et al. (2018) investigated both the conventional Cuttlefish Algorithm (CFA) and its chaotic variant for IDS feature selection, reporting detection rates of 92.55% and 96.01%, respectively. Similarly, Saheed (2022) employed a Binary Firefly Algorithm for feature optimization and achieved a detection rate of 99.72%. More recently, Sandhya et al. (2024) developed a hybrid IDS framework that integrated ACO with a deep learning model for feature selection and classification, achieving a detection rate of 99.87%.

Although these studies highlight the effectiveness of evolutionary and hybrid metaheuristic approaches for IDS feature optimization, their reported performances remain marginally lower than that of the proposed ACO-CFA framework. The superior detection rate achieved by the proposed model can be attributed to the complementary strengths of ACO's global search capability and CFA's local optimization mechanism, which collectively facilitate the selection of highly discriminative and non-redundant feature subsets for classification. Overall, despite differences in datasets, model architectures, optimization strategies, and experimental settings across the reviewed studies, the proposed hybrid ACO-CFA framework consistently demonstrates enhanced detection capability and competitive performance. The comparative findings therefore validate the effectiveness, robustness, and practical applicability of the proposed framework for intrusion detection tasks in modern cybersecurity environments.

Table 6. Performance comparison of the proposed feature selection framework against the State-of-the-Art in IDS.

Author(s)	Frameworks	Accuracy (%)
Aghdam & Kabiri (2016)	ACO	98.90
Balasaraswathi et al. (2018)	CFA	92.55
Balasaraswathi et al. (2018)	Chaotic CFA	96.01
Saheed (2022)	Binary Firefly Algorithm	99.72
Sandhya et al. (2024)	Hybrid ACO+DELM	99.87
Proposed	Hybrid ACO-CFA	99.97

CONCLUSION

This study successfully proposed and systematically evaluated a novel hybrid feature selection framework for Intrusion Detection Systems (IDS), integrating the principles of Ant Colony Optimization (ACO) and the Cuttlefish Algorithm (CFA). The findings demonstrated that the proposed bio-inspired hybrid approach substantially improves IDS performance through the effective identification and selection of relevant, non-redundant features from complex and high-dimensional network traffic datasets. By addressing the limitations associated with conventional and static feature selection techniques, the proposed framework enhances the efficiency and reliability of intrusion detection in cybersecurity environments. Experimental evaluations conducted across multiple training-to-valiation/testing partition ratios revealed consistently high detection performance, with the selected classifiers, Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF), achieving excellent predictive accuracy and near-perfect ROC–AUC values.

The empirical results further highlight the complementary strengths of integrating ACO's global search capability with CFA's localized optimization process, thereby enabling effective feature refinement and improved classification performance while reducing computational complexity. Among the evaluated classifiers, the RF model emerged as the most robust and suitable for practical implementation due to its superior accuracy and resilience, whereas the DT model provided a reliable alternative in scenarios emphasizing computational efficiency and interpretability. Overall, this research contributes significantly to the field of cybersecurity by introducing a scalable, efficient, and high-performing IDS framework capable of strengthening defenses against evolving cyber threats and improving the adaptability and effectiveness of intrusion detection mechanisms in real-world applications. Future research will focus on the comprehensive application of the proposed hybrid ACO–CFA feature selection framework across multiple benchmark datasets, including NSL-KDD, CIC-IDS2017, and UNSW-NB15, in addition to the KDD Cup 99 dataset used in this study. Furthermore, a comparative performance analysis will be conducted to evaluate and contrast the results obtained across these datasets with the findings reported in the present work.

REFERENCE

Aghdam, M. H., & Kabiri, P. (2016). Feature Selection for Intrusion Detection System Using Ant Colony Optimization. *International Journal of Network Security 18*(3), 420--432.

Akoul, N., Ahmad, A. A., & Pro, S. (2026). A hybrid approach for network intrusion detection using artificial bee colony optimization and ensemble learning. *Journal of Engineering and Applied Science, 73*(37), 1--22.

Balasaraswathi, V. R., Sugumaran, M., & Hamid, Y. (2017). Feature selection techniques for intrusion detection using non-bio-inspired and bio-inspired optimization algorithms. *Journal of Communication and Information Networks, 2*(4), 107--119.

Balasaraswathi, V. R., Sugumaran, M., & Hamid, Y. (2018). Chaotic Cuttle Fish Algorithm for Feature Selection of Intrusion Detection System. *International Journal of Pure and Applied Mathematics*, *119*(10), 921--935.

Balasaraswathi, V. R., & Sugumaran, M. (2019). A Hybrid Algorithm Using Ant Colony Optimisation and Cuttle Fish Algorithm for Feature Selection of Intrusion Detection. *International Journal of Scientific & Engineering Research*, *10*(1), 1807--1814. http://www.ijser.org

Eesa, A. S., Orman, Z., & Brifcani, A. M. A. (2015). A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. *Expert Systems with Applications*, *42*(5), 2670--2679. https://doi.org/10.1016/j.eswa.2014.11.009

Eesa, A.S., Sadiq, S., Hassan, M. M., & Orman, Z. (2021). Rule generation based on modified Cuttlefish Algorithm for Intrusion Detection System. *Uludağ University Journal of The Faculty of Engineering*, 26(1), 253--267. https://doi.org/10.17482/uumfd.747078

Kamil, S., Sheikh Abdullah, S.N.H., Firdaus, A. and Usman, O.L. (2022), The Rise of Ransomware: A Review of Attacks, Detection Techniques, and Future Challenges. Proceedings of the *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. Dubai, United Arab Emirates, 16-17 February 2022. DOI: 10.1109/ICBATS54253.2022.9759000

Kannan, A. V. (2020). Intrusion Detection in Internet of Things using Ant Colony Optimisation. *ICTACT Journal on Data Science and Machine Learning*, 1(3), 88--91.

Ogundokun, R. O., Misra, S., Bajeh, A. O., Okoro, U. O., & Ahuja, R. (2022). *An Integrated IDS using ICA-based Feature Selection and SVM Classification Method* (S. Misra & C. Arumugam (eds.)). Springer Nature.

Panliang, M., Madaan, S., Ahmed, S., Ali, B., Gowrishankar, J., & Khatibi, A. (2025). Enhancing feature selection for multi-pose facial expression recognition using a hybrid of quantum inspired firefly algorithm and artificial bee colony algorithm. *Scientific*

Reports, 15(4665), 1--23.

Prithi, S., & Sumathi, S. (2024). A technical research survey on bio-inspired intelligent optimization grouping algorithms for finite state automata in intrusion detection system. MultiCraft International *Journal of Engineering, Science and Technology, 16*(2), 48--67. https://doi.org/10.4314/ijest.v16i2.6

Radhakrishnan, S., Aljawarneh, P., & Kumar, V. (2019). Snort – Lightweight intrusion detection for networks. In Proceedings of the *13th USENIX Conference on System Administration (LISA '99),* 229--238.

Rufai, K. I., Usman, O. L., Olusanya, O. O. and Adedeji, O. B. (2021), Solving Travelling Salesman Problem using an Improved Ant Colony Optimization Algorithm. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR), 6*(2): 158--170.

Rufai, K.I., Muniyandi, R. C., and Usman, O.L. (2016), Tacking the Course of Dimensionality in Intrusion Detection Systems: Membrane Computing Approach. Proceedings of the 2nd TASUED-UCC International Conference, Tai Solarin University of Education, Nigeria, August 22nd -25th, 2016, pp. 1539--1549.

Safana, I. Y., Obunadike, G. N. & Surajo, Y. (2026). An Explainable Ensemble Deep Learning Approach for Intrusion Detection in Industrial Internet of Things. *Journal of Basics and Applied Sciences Research, 4*(1): 143--151.

Saheed, Y. K. (2022). *A Binary Algorithm based Feature Selection Method on High Dimensional Intrusion Detection Data* (S. Misra & C. Arumugam (eds.)). Springer Nature.

Sandhya, E., Benschwartz, R., Sathiya, T., Sangeetha, M., Sreeramamurthy, K., & Preetha, M. (2024). Hybrid Ant Colony Optimization and Deep Learning for Anomaly Intrusion Detection. *International Journal of Intelligent Systems and Applications in Engineering, 12*(19), 873--881. www.ijisae.org

Usman, O.L., Muniyandi, R.C., Omar,K., Mohmad, M., Owoade, A.A. & Kareem, M.A. (2025). HoRNS-CNN Model: An Energy-Efficient Fully Homomorphic Residue Number System Convolutional Neural Network Model for Privacy-Preserving Classification of Dyslexia Neural-Biomarkers. *Brain Informatics, Springer Nature, 12*(11), 1--28. https://doi.org/10.1186/s40708-025-00256-z

Usman, O. L. (2025). Identifying Significant Structural Factors associated with Knee Pain Severity in Patients with Osteoarthritis using Hybrid Bio-BERT Bi-LSTM CNN Model. *Journal of Science and Information Technology (JOSIT), 19*(2): 18--28.

Usman O.L, & Adeusi O.O. (2025). Optimization of an efficient net-based transfer learning model for automated pneumonia detection from chest x-ray images. *Dutse Journal of Pure and Applied Sciences, 11*(4a), 397--411. https://www.ajol.info/index.php/dujopas/article/view/311486

Usman, O. L., Adeusi, O. O., Kareem, M. A., Owoade, A. A. & Muniyandi, R. C. (2026). Quantitative study on impact of EfficientNet-based deep transfer learning model for pneumonia detection with explainable artificial intelligence using chest radiographs. *Zamfara International Journal of Science, Technology, Education and Mathematics. 3*(1), 45--56. https://doi.org/10.64348/zije.2026344

Usman, O. L. & Muniyandi, R. C. (2020). CryptoDL: Predicting Dyslexia Biomarkers from Encrypted Neuroimaging Dataset Using Energy-Efficient Residue Number System and Deep Convolutional Neural Network. *Symmetry MDPI-Basel, 12*(836), 1--24. https://doi.org/10.3390/sym12050836

Usman, O. L., Muniyandi, R. C., Omar, K., & Mohamad, M. (2021). Advanced machine learning methods for dyslexia biomarker detection: A review of implementation details and challenges. *IEEE Access, 9*, 36879–36894. https://doi.org/10.1109/ACCESS.2021.3062709

Yinka-Banjo, C., Alli, P., Misra, S., Oluranti, J., & Ahuja, R. (2022*). Intrusion Detection using Anomaly Detection Algorithm and Snort* (S. Misra & C. Arumugam (eds.)). Springer Nature.

Zivkovic, M., Tair, M., Venkatachalam, K., Bacanin, N., & Trojovský, P. (2022). Novel hybrid firefly algorithm: an application to enhance XGBoost tuning for intrusion detection classification. *PeeJ Computer Science, 8*(e956), 1--38. https://doi.org/10.7717/peerj-cs.956