

Grey-Lag Goose Algorithm for Feature Selection in DDoS Attack Detection: A Novel Application on CICIDS2019 Dataset

Oladunjoye, John Abiodun¹; Andrew Ishaku Wreford^{2*}; Giliki, Jerison³

¹Department of Computer Science, Faculty of Computing and Information Systems, Federal University Wukari, Wukari, Nigeria

²Department of Software Engineering, Faculty of Computing and Information Systems, Federal University Wukari, Wukari, Nigeria

³Department of Computer Science, Adamawa State College of Education, Hong, Adamawa State, Nigeria.

*Corresponding Author

Andrew Ishaku Wreford

Article History

Received: 11.07.2026

Revised: 20.08.2026

Accepted: 07.09.2026

Published: 28.09.2026

Abstract: Distributed Denial of Service (DDoS) attacks are still one of the focal cybersecurity concerns, primarily due to their complex nature and the sheer number of devices attached. Traditional detection lags in keeping up with emerging attack behaviors, and machine learning-based models have difficulty handling high-dimensional data, like the CICIDS2019 dataset with over 80 features. Feature selection (FS) has thus emerged as an essential preprocessing step to boost detection performance, decrease the cost of computation, and facilitate improved interpretation of the model. In this current study, the GreyLag Goose Algorithm (GGA), a bio-inspired metaheuristic optimization algorithm derived from greylag geese's migratory and foraging behavior, is utilized for the first time to solve the problem of feature selection in DDoS detection. GGA's balanced exploration-exploitation allows effective exploration of high-dimensional spaces of features to find the most informative subset of features. The extracted features were evaluated using the CICIDS2019 dataset with Support Vector Machine (SVM) and Random Forest (RF) classifiers. Experimental results show that GGA reduced the dimensionality from over 80 to 29 features with better classification performance. RF achieved closely perfect with 99.95% accuracy and AUC of 0.9999, although slightly better than SVM's 99.47% accuracy and 0.9987 AUC value. The findings underscore the strength of GGA in feature selection optimization of DDoS detection to achieve an efficient and stable approach that is a balance between efficacy and high prediction accuracy.

Keywords: GreyLag Goose Algorithm (GGA), Bio-inspired Optimization, Metaheuristic Algorithms, Feature Selection (FS), Distributed Denial of Service (DDoS).

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

Cite this article:

Oladunjoye, J. A.; Wreford, A. I.; Giliki, J., (2026). Grey-Lag Goose Algorithm for Feature Selection in DDoS Attack Detection: A Novel Application on CICIDS2019 Dataset. *ISAR Journal of Multidisciplinary Research and Studies*, 4(9), 31-39.

1.0 Introduction

Distributed Denial of Service (DDoS) attacks are among the most persistent and devastating network infrastructure threats, causing tremendous financial losses and disruptions to service in many industries. The sheer magnitude of internet-of-things devices coupled with the sophisticated nature of the attack vectors have made DDoS attacks more common and complex, necessitating for evolved detection tools that can keep pace with shifting threat profiles [1][2]. Traditional signature-based methods often fail to recognize novel attack patterns, while machine learning approaches are plagued with the problem of high-dimensional feature spaces and computational complexity [3][4].

The performance of machine learning-driven DDoS detection systems is heavily dependent on the quality and suitability of selected features of network traffic data. High-dimensional datasets, such as the CICIDS2019 dataset with over 80 features, present significant challenges including the curse of dimensionality, increased computational overhead, and potential overfitting [5][6]. Feature selection becomes a critical preprocessing task that can significantly enhance the detection accuracy at the cost of lower computational cost and increased interpretability of the model [7][8].

Recent trends in cyber security research have shown the significance of feature selection to maximize intrusion detection system performance. Traditional methods for selecting features, such as filter-based methods like correlation-based feature selection and wrapper-based methods like recursive feature elimination, have been effective but often suffer from limitations in handling complex, non-linear relationships in network traffic data [9][10]. Embedded techniques, though computationally useful, do not always capture the global optimal feature subset because they are subject to the particular learning algorithm used [11].

The emergence of metaheuristic optimization algorithms has opened up new possibilities for feature selection in cybersecurity systems. Genetic algorithm, particle swarm optimization, and ant colony optimization have been found useful in feature selection problems in network intrusion detection and proved better than traditional methods [12][13][14]. The algorithms, however, suffer from the drawbacks of premature convergence, local optima trap, and heavy computational complexity, particularly while dealing with high-dimensional cybersecurity data [15][16].

Bio-inspired optimization algorithms have found extensive applications in the area of cybersecurity due to their high potential for solving hard, multi-modal optimization problems with efficiency. Bio-inspired algorithms rooted in natural phenomena and biological systems deliver robust solutions for parameter tuning, feature selection, and pattern

recognition tasks in network security solutions [17][18]. Recent studies reveal that bio-inspired algorithms can effectively balance exploration and exploitation, thus offering improved convergence characteristics and better global optima detection [19][20].

Swarm intelligence techniques like Grey Wolf Optimizer, Whale Optimization Algorithm, and Firefly Algorithm have been used successfully in solving most of the cybersecurity problems, from malware detection to intrusion detection and vulnerability scoring [21][22][23]. The algorithms are particularly effective in handling high-dimensional feature spaces and non-linear optimization landscapes that are normally encountered in cybersecurity problems [24][25].

The GreyLag Goose Algorithm (GGA), relatively recently introduced as a novel bio-inspired optimization algorithm, is inspired by the social behavior and migratory characteristics of greylag geese. The algorithm is characterized by its peculiar balance between exploration phase and exploitation phase, which is inspired by the foraging pattern, flocking, and migratory behaviors of the geese [26][27]. GGA features several distinct characteristics like adaptive control of step size, handling population dynamically, and multi-objective optimization capabilities which render it extremely suitable for challenging optimization problems [28].

Mathematically, GGA is founded on three distinct behavioral phases: foraging behavior for local search, flocking behavior for information sharing, and migration behavior for global searching. The algorithm employs a position update mechanism combining individual experience, social learning, and environmental adaptation with improved convergence speed and solution quality compared to traditional optimization algorithms [29][30]. Recent comparative studies have shown GGA's stronger performance on various optimization benchmarks, particularly for high-dimensional and multi-modal optimization problems [31][32].

Despite the promising performance of bio-inspired algorithms in cybersecurity operations, the application of the GreyLag Goose Algorithm to feature selection towards the detection of DDoS attacks has never been tried, thus this study to explore its potentialities since existing feature selection approaches to DDoS detection either base themselves on traditional optimization methods or proven bio-inspired algorithms, and there is a wide research gap on exploring the potential of newer and more sophisticated bio-inspired algorithms like GGA [33][34].

2.0 Related Work

The paper [35] introduced an Improved Whale Optimization Algorithm (IWOA) to be used for feature selection and an Optimized Long Short-Term Memory (OLSTM) for classification. The proposed solution achieved a high rate of

detection with accuracy at 97.12%, precision at 96.74%, recall at 96.27%, and F1-score at 96.54% on the CIC-DDoS 2019 dataset. This method outperformed others like CNN-BiLSTM and CNN-LSTM, citing the ability of IWOA to enhance classification by determining good features. A new algorithm for detecting Distributed Denial of Service (DDoS) attacks in Software-Defined Networks (SDNs) was created by [36] using Grey Wolf Optimizer (GWO) as a feature selector and a Support Vector Machine (SVM) classifier. Their work identifies the deficiency of existing SDN-based DDoS detection methods in low precision and high false negative rates caused by inferior feature extraction. To address this, they employed a binary Grey Wolf optimization algorithm on the InSDN dataset with notable improvement in detection. The unimodal equation achieved a 97% F1-score, while the multi-modal equation achieved 100% accuracy, 96% recall, and a 98% F1-score. [37] proposed a hybrid detection framework to combat Distributed Denial of Service (DDoS) attacks in healthcare networks, particularly against the vulnerabilities introduced by the Internet of Medical Things (IoMT). In their work, they integrate a deep neural network architecture that combines CNN and LSTM with Particle Swarm Optimization (PSO) for better training efficiency and detection performance. Utilizing the CIC-IDS-2019 dataset, the framework achieved a detection accuracy of 99.3%, outperforming some conventional algorithms.

[38] constructed an effective Distributed Denial-of-Service (DDoS) intrusion detection system using a Convolutional Neural Network (CNN) together with an Inception mechanism to enable better classification of all types of attacks. The authors addressed the primary challenges encountered when detecting DDoS, including imbalanced data, difficulty in multi-classification, and computational expense, by introducing an innovative feature selection and preprocessing method. On the CICDDoS2019 benchmarking dataset, the model outperformed baseline and state-of-the-art approaches with 96.82% accuracy, 96.82% recall, 96.76% precision, and F1-score of 96.50%. [39] proposed a hybrid intrusion detection system (IDS) that identifies DoS attacks using stacked feature selection techniques and ensemble learning classifiers. To break through the challenges of high-dimensional information and the lack of strength of common base classifiers, the study employed a novel feature selection technique combining Random Boruta Selector (RBS), Relief, and Pearson Correlation Coefficient (PCE) to identify optimal feature sets. The extracted features were also classified by stacked learning classifiers to enhance the detection accuracy. The proposed model was evaluated on the CICDDoS-2019 dataset with 96.5% accuracy, 96.0% precision, 95.8% recall, and a 96% F1-score, which is better than traditional classifiers in identifying DoS attacks.

3.0 Research Methodology

This study adopts a quantitative approach using the GGA, a nature-inspired metaheuristic, for feature selection in detecting DDoS attacks on the CICDDoS2019 dataset. The process begins with loading and preprocessing the dataset through one-hot encoding and min-max normalization. GGA then generates candidate feature subsets, designating the best as leader, while others align and update positions through migration. Leadership is rotated periodically to balance exploration and exploitation, and each subset is evaluated using a fitness function until the stopping criterion is met. The selected features are then used to train SVM and RF models, whose performance is assessed; if unsatisfactory, the process restarts from preprocessing. The overall workflow is illustrated in Figure 1.

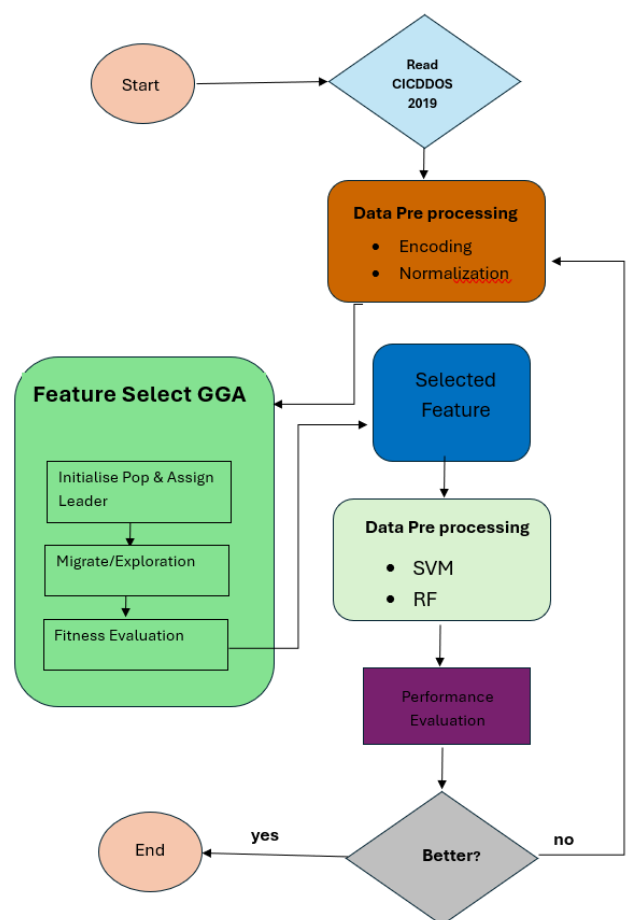


Figure 1: Research Method

3.1 Dataset Description

The dataset used for the DDoS attack detection is CIC-DDoS 2019 [40] dataset which contains 86 features of network traffic package generated by open-source tool. It created network packets and collects features from them. The DDoS attacks were based on the use of reflection against authorized servers like Domain Name Server (DNS), Network Basic Input/Output System (NETBIOS) and Simple Network Management Protocol (SNMP) and

Lightweight Directory Access Protocol (LDAP) that lower many services across network.

3.2 Preprocessing

The CIC-DDoS 2019 dataset undergoes two well-known preprocessing steps namely normalization and one-hot encoding. Both of these steps prepare the raw network traffic data for effective machine learning processing. During feature encoding, symbolic attributes of the dataset are encoded to numerical representations through one-hot encoding. The process translates categorical variables into binary vectors with the aim of having features represented numerically. We also apply min-max normalization to normalize every numerical feature to the same range [0, 1]. This transformation reduces the effect of outliers and minimizes variance among different traffic features. The principle is to give equal feature weighing when training the model. Normalization was performed through:

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

Where X'_{ij} is the normalized value of feature j in sample i , X_{ij} is the original value of j in sample i , $\min(X_j)$ is the minimum value of feature j across all samples and $\max(X_j)$ is the maximum value of feature j across all samples.

3.3 Feature Selection

The GGA is inspired by Greylag geese's migratory patterns, which optimally position themselves in V-shaped formations to minimize energy expenditure during long-distance flights [41][42]. Analogously, in feature subset selection, every goose is a candidate feature subset and the V-shape is flying together towards best solutions. Exploration (searching new regions in the feature space) and exploitation (focusing on promising solutions) are exploited by the algorithm to use both in controlling the optimization process.

Mathematical Formulation

Let the dataset be denoted as

$$D = \{(X_i, Y_i) | i = 1, 2, \dots, N\} \quad (2)$$

where $X_i \in \mathbb{R}^d$ represents the d -dimensional feature vector, and $Y_i \in \{0, 1\}$ denotes the class label (benign or attack). The goal is to select an optimal subset of features

$F^* \subseteq \{1, 2, \dots, d\}$ such that the classification performance is maximized.

Each goose in the population encodes a binary solution vector:

$$G_j = (g_{j1}, g_{j2}, \dots, g_{jd}), g_{jk} \in \{0, 1\} \quad (3)$$

where $g_{jk} = 1$ indicates that feature k is selected by goose j .

The fitness of each goose is evaluated using a hybrid objective function that balances classification accuracy and feature reduction. This is mathematically expressed as:

$$f(G_j) = \alpha \cdot \text{Acc}(G_j) - \beta \cdot |G_j|/d \quad (4)$$

where $\text{Acc}(G_j)$ denotes the classification accuracy obtained using the subset represented by G_j , $|G_j|$ is the number of selected features, and α and β are weighting coefficients controlling the trade-off between accuracy and feature reduction.

The movement of geese in GLGA is governed by the leader-following mechanism. The leader goose represents the current best solution G^* . Other geese update their positions by aligning with the leader while introducing perturbations to maintain diversity. The position update can be expressed as:

$$G_j^t + 1 = \text{Sigmoid}(G_j^t + \lambda \cdot (G^* - G_j^t) + \eta \cdot \epsilon) \quad (5)$$

where t denotes the iteration, λ is the alignment factor, η controls randomness, ϵ is a stochastic perturbation term, and the Sigmoid function is applied to map continuous values into binary space.

For CICDDoS2019 dataset preprocessing, optimization is carried out afterwards. Cleaning of the dataset is performed initially by handling missing values and normalizing continuous features using Min-Max normalization. The GGA population is then initialized with candidate feature subsets, where each is a binary selection vector over the 80+ available traffic features in the dataset. The feature subsets are assessed by a machine learning classifier such as Random Forest (RF) and Support Vector Machine (SVM). Each feature subset selected by a goose is used to train and test the classifier through k-fold cross-validation in order to

give a reliable estimation of classification accuracy. The fitness function gives a balance between the number of features to be minimized and classification performance to be maximized.

Throughout the iterative process, the leader goose represents the optimal subset of features. Other geese refine their subsets through imitation of the leader, and stochastic perturbations maintain diversity and prevent premature convergence. Over iterations, the algorithm converges to a

compact subset of features that capture the discriminative patterns within network traffic.

Finally, the optimal feature subset F^* identified by GLGA is used to retrain the classifier on the condensed data, which achieves higher computational efficiency with no sacrifice, and even with gain, in accuracy in classifying benign vs. attack traffic. In addition to reducing dimensionality, this method also enhances interpretability by highlighting the most informative features involved in DDoS detection.

Algorithm 1: GreyLag Goose Algorithm for Feature Selection

Input: CICDDoS2019 dataset with d features

Output: Optimal feature subset F^*

1. Initialize population of geese P with random binary vectors of length d
 2. Evaluate fitness $f(G_j)$ for each goose using classifier accuracy and feature ratio
 3. Identify leader goose G^* with the best fitness
 4. While stopping criterion not met (max iterations or convergence):
 - For each goose G_j in population:
 - Update position:

$$G_j = \text{Sigmoid}(G_j + \lambda * (G^* - G_j) + \eta * \epsilon)$$
 - Evaluate new fitness $f(G_j)$
 - Update leader G^* if a better goose is found
 5. Return feature subset corresponding to G^*
-

4.0 Result Discussion

4.1 Environmental Setup

The DDOS attack detection model was developed using the Anaconda computational environment. The experiment was conducted on a Windows operating system with a dual-core Intel Core i5 CPU and 8 GB of RAM. The implementation was carried out in Python using essential machine learning libraries such as Scikit-Learn, NumPy, and pandas. The

visualization of model performance was achieved using Matplotlib and Seaborn.

4.2 Feature Selection Result

As aforementioned, GGA was proposed for FS where the goal is to choose the most informative subset of features while balancing model accuracy and complexity. Each parameter in the table 1 controls an aspect of this optimization process.

Table 1: GGA Feature Selection Parameters

Parameter	Type	Default Value	Description
base_estimator	BaseEstimator	Required	The base ML model (e.g., Logistic Regression, SVM, etc.) to evaluate feature subsets.
Alpha	float	0.9	Weight controlling the contribution of accuracy in the fitness function.
Beta	float	0.1	Weight controlling the penalty for selecting too many features.
pop_size	int	24	Number of geese (population size) in the flock.
Iters	int	25	Number of iterations for optimization.
align_lambda	float	0.9	Degree of alignment towards the leader goose (controls exploitation).
Eta	float	0.2	Scale of Gaussian noise (controls exploration/diversity).
cv_folds	int	3	Number of cross-validation folds to evaluate fitness.
random_state	int	42	Random seed for reproducibility of results.

In summary, these parameters strike a balance between accuracy ($\alpha = 0.9$) and parsimony ($\beta = 0.1$), with enough exploration ($\eta = 0.2$) and exploitation ($\lambda = 0.9$). The design reflects a preference for high model accuracy while avoiding excessive feature sets.

Thus, the parameter of GGA from table 1 results to the selected feature from table 2. The features initially have 80 features, out of which 29 features were selected.

Table 2: Selected Feature

Feature	Feature	Feature
Total Backward Packets	Bwd IAT Max	Avg Fwd Segment
Fwd Packets Length Total	Fwd PSH Flags	Avg Bwd Segment Size
Fwd Packet Length Max	Bwd PSH Flags	Fwd Avg Bulk Rate
Fwd Packet Length Mean	Fwd URG Flags	Bwd Avg Packets/Bulk
Bwd Packet Length Max	Bwd URG Flags	Subflow Bwd Packets
Flow Bytes/s	Bwd Header Length	Subflow Bwd Bytes
Fwd IAT Total	Packet Length Mean	Init Fwd Win Bytes
Idle Min	FIN Flag Count	Init Bwd Win Bytes
Fwd IAT Min	CWE Flag Count	Idle Max
Bwd IAT Total	Avg Packet Size	Idle Min

Essentially, the feature selection process reduced the dimensionality of the dataset to a smaller subset of highly informative attributes. Key points to note are:

- i. Traffic Flow Features:
 - a. Total Backward Packets, Fwd Packets Length Total, Flow Bytes/s captures data transfer volume and direction and are crucial in distinguishing normal vs. attack traffic.
- ii. Packet Length and Timing Features:
 - a. Fwd Packet Length Max/Mean, Bwd Packet Length Max, Packet Length Mean, Avg Packet Size describes variability in packet sizes. Affirming to the facts that attack traffic often exhibits abnormal distributions.
 - b. Fwd IAT Total, Bwd IAT Total, Idle Min/Max defines the inter-arrival times and idle periods capture timing anomalies. Importantly, the DDoS or flooding attacks may show reduced idle times and uniform IATs.
- iii. TCP Flag Features:
 - a. Fwd PSH Flags, Bwd PSH Flags, Fwd URG Flags, Bwd URG Flags, FIN Flag Count, CWE Flag Count are abnormal flag combinations often indicate suspicious sessions (e.g., port scans, connection resets).

- iv. Flow Behavior Features:

- a. Fwd Avg Bulk Rate, Bwd Avg Packets/Bulk, Subflow Bwd Packets/Bytes reflects session-level statistics that capture volumetric or bursty attack behavior which are also important for attack detection.

4.3 Result Presentations

Table 3 and figure 2 present the result of the RF and SVM algorithm on the selected features for attack detection from table 2. Metrics considered include accuracy, precision, recall, f1-score, and area under curve.

Table 3: Result Presentation

Algorithm	Accuracy	Precision	Recall	F1-score	AUC
RF	0.9995	0.9998	0.9995	0.9997	0.9999
SVM	0.9947	0.9975	0.9956	0.9966	0.9987

The comparison of the performance of the attack detection models demonstrates that both the RF and SVM algorithms exhibited great performance in all the performance metrics, indicating their viability for intrusion detection.

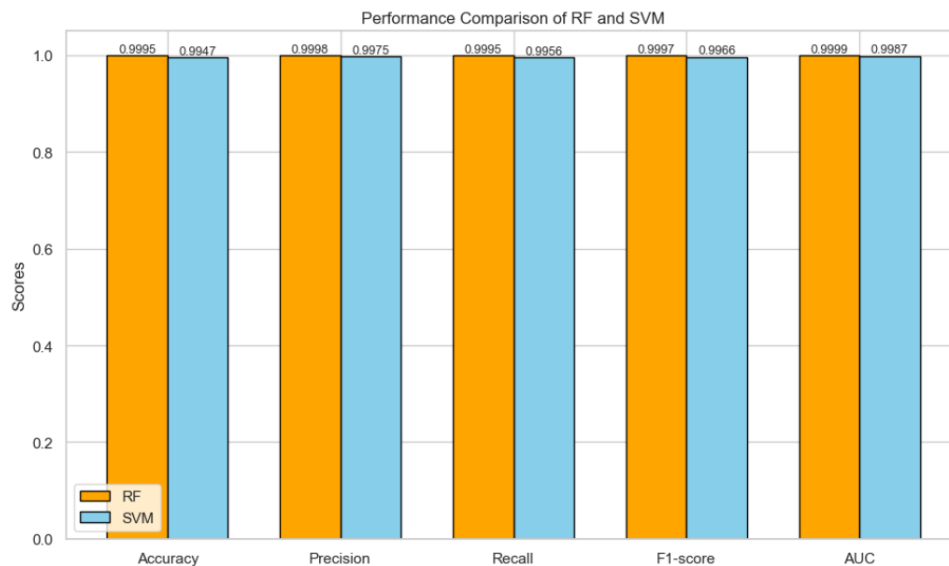


Figure 2: Result Comparison

The RF did slightly better than the SVM in all metrics. It was 99.95% accurate for predicting that nearly all the attack and normal traffic samples were classified correctly. With accuracy of 99.98%, it seldom labeled benign traffic as an attack, thereby minimizing false alarms a primary requirement in real-world security systems. Moreover, its 99.95% recall guarantees that almost all cases of attacks were correctly identified, with barely any cases of missed threats. Its phenomenally high F1-score of 99.97% tells us that the model had a very high precision and recall ratio, while the 0.9999 AUC score tells us that it has nearly perfect ability to distinguish between normal and attack traffic.

SVM, however, was equally robust. It was 99.47% accurate with an AUC of 0.9987 and was highly resistant to discriminating between attacks and normal traffic. Its accuracy (99.56%) and recall (99.75%) indicate SVM was highly accurate in eliminating false positives and capturing attacks, although slightly less reliable than RF. Its F1-score (99.66%) is also an indication of its ability for a good trade-off between precision and recall, and it is also a good candidate to detect attacks.

Overall, while both algorithms were excellent, RF had a slight lead over SVM and is the more trustworthy option to use in actual intrusion detection systems. The results show ensemble-based strategies like RF, by making decisions of many decision trees and combining them, are a right option to manage complexity and diversity in attack traffic. But the consistently high performance of SVM confirms that kernel-based classifiers remain viable choices, especially in cases where interpretability or computational efficiency is needed.

Finally, the GGA parameter setting was essential in model optimization. With a proper balance between feature reduction and accuracy, GGA chose useful traffic-based features capturing volume, timing, and flag-level behaviors

of importance for intrusion detection. The optimization, on the other hand, witnessed RF, particularly, achieve perfect detection with the subset of features reduced, thereby confirming the effectiveness of GGA in model efficiency and accuracy improvement.

5.0 Conclusion

In this paper, the GGA was introduced as a novel feature selection algorithm for DDoS attack detection on the CICIDS2019 dataset. As opposed to conventional filter- and wrapper-based feature selection approaches, GGA leverages bio-inspired migratory and flocking behaviors to optimally address the tradeoff between global exploration and local exploitation and thereby address the problems of high-dimensional network traffic data. Experimental findings confirmed that GGA effectively projected the initial feature space to a lower-dimensional subset of highly informative 29 features effectively, lowering computational complexity significantly while increasing classification accuracy. SVM and RF classifiers learned using the optimized feature set enhanced detection, with RF achieving 99.95% accuracy and nearly ideal AUC measure of 0.9999. These results confirm GGA's ability not only to remove feature redundancy but also enhance the discriminant abilities of the detection models. The process of optimization also showed essential traffic features like flow behaviors, packet length distributions, and TCP flag anomalies that pointed toward their importance in classifying benign versus malicious traffic behaviors. Overall, this paper demonstrates that GGA is an effective metaheuristic to use in tackling cybersecurity issues such as intrusion detection system feature selection. Its ability to tackle the CICIDS2019 dataset demonstrates its potential to generalize to other cyber-threat detection tasks, therefore making it a suitable option to be part of the family of bio-inspired optimization algorithms for network security.

Follow-up studies can continue this work by integrating GGA into deep learning-based intrusion detection systems or applying it to real-time DDoS detection in big network scenarios.

References

- Mishra, N., & Pandya, S. (2021). Internet of things applications, security challenges, attacks, intrusion detection, and future visions: A systematic review. *IEEE Access*, 9, 59353-59377.
- Malhotra, P., Singh, Y., Anand, P., Bangotra, D. K., Singh, P. K., & Hong, W. C. (2021). Internet of things: Evolution, concerns and security challenges. *Sensors*, 21(5), 1809.
- Anikolova, E., Martins, S., Rozental, D., Fontana, J., & Maier, P. (2024). Ransomware detection through behavioral attack signatures evaluation: A novel machine learning framework for improved accuracy and robustness. *Authorea Preprints*.
- Ahmed, U., Nazir, M., Sarwar, A., Ali, T., Aggoune, E. H. M., Shahzad, T., & Khan, M. A. (2025). Signature-based intrusion detection using machine learning and deep learning approaches empowered with fuzzy clustering. *Scientific Reports*, 15(1), 1726.
- Talukder, M. A., Islam, M. M., Uddin, M. A., Hasan, K. F., Sharmin, S., Alyami, S. A., & Moni, M. A. (2024). Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction. *Journal of big data*, 11(1), 33.
- Roopak, M., Parkinson, S., Tian, G. Y., Ran, Y., Khan, S., & Chandrasekaran, B. (2024). An unsupervised approach for the detection of zero-day distributed denial of service attacks in Internet of Things networks. *IET Networks*, 13(5-6), 513-527.
- Cheng, X. (2024). A comprehensive study of feature selection techniques in machine learning models. *Available at SSRN 5154947*.
- Dhanka, S., Sharma, A., Kumar, A., Maini, S., & Vundavilli, H. (2025). Advancements in Hybrid Machine Learning Models for Biomedical Disease Classification Using Integration of Hyperparameter-Tuning and Feature Selection Methodologies: A Comprehensive Review. *Archives of Computational Methods in Engineering*, 1-36.
- Theng, D., & Bhoyar, K. K. (2024). Feature selection techniques for machine learning: a survey of more than two decades of research. *Knowledge and Information Systems*, 66(3), 1575-1637.
- Amiribrahimabadi, M., & Mansouri, N. (2024). A comprehensive survey of feature selection techniques based on whale optimization algorithm. *Multimedia Tools and Applications*, 83(16), 47775-47846.
- Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiaq, T., Rafa, N., ... & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56(11), 13521-13617.
- Fayadh, S. M. (2025). Hybrid Machine Learning Model for Feature Selection in DDoS Attack Detection in Cloud Environments Using Convolutional Neural Networks and Genetic Algorithms. *Wasit Journal for Pure sciences*, 4(1), 94-103.
- Saeed, A. A., & Jameel, N. G. M. (2021). Intelligent feature selection using particle swarm optimization algorithm with a decision tree for DDoS attack detection. *International Journal of Advances in Intelligent Informatics*, 7(1), 37-48.
- Saeed, A. A., & Jameel, N. G. M. (2021). Intelligent feature selection using particle swarm optimization algorithm with a decision tree for DDoS attack detection. *International Journal of Advances in Intelligent Informatics*, 7(1), 37-48.
- Xu, T., Ma, C., Su, Z., Su, J., Ma, Z., Wang, J., & Yang, Z. (2025). Enhancing Cybersecurity in the Big Data Era: A GA-Optimized Fuzzy Clustering Approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 33(03), 353-377.
- Chui, K. T., Liu, R. W., Zhao, M., & Zhang, X. (2024). Bio-inspired algorithms for cybersecurity-a review of the state-of-the-art and challenges. *International Journal of Bio-Inspired Computation*, 23(1), 1-15.
- Pham, T. H., & Raahemi, B. (2023). Bio-inspired feature selection algorithms with their applications: a systematic literature review. *IEEE Access*, 11, 43733-43758.
- Saadouni, R., Gherbi, C., Aliouat, Z., Harbi, Y., & Khacha, A. (2024). Intrusion detection systems for IoT based on bio-inspired and machine learning techniques: a systematic review of the literature. *Cluster Computing*, 27(7), 8655-8681.
- Chou, J. S., & Molla, A. (2022). Recent advances in use of bio-inspired jellyfish search algorithm for solving optimization problems. *Scientific Reports*, 12(1), 19157.
- Devika, G., & Karegowda, A. G. (2023). Bio-inspired optimization: algorithm, analysis and scope of application. In *Swarm Intelligence-Recent Advances and Current Applications*. IntechOpen.
- Prashanth, S. K., Iqbal, H., & Illuri, B. (2023). An enhanced grey wolf optimisation-deterministic convolutional neural network (GWO-DCNN) model-based IDS in MANET. *Journal of Information & Knowledge Management*, 22(04), 2350010.
- Vadakkethil, S. E., Polimetla, K., Velpula, S., Pareek, P. K., & Sontakke, D. (2024, April). Improved Whale Optimization Algorithm and Optimized Long Short-

- Term Memory for DDoS Cyber Security Threat. In *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* (pp. 01-05). IEEE.
23. Suggala, M., Hari, S., & Mallampati, S. B. (2024, March). Firefly-Driven Feature Selection for Enhanced DDoS Detection in IoT Networks. In *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)* (pp. 1-6). IEEE.
 24. Asghari, K., Masdari, M., Gharehchopogh, F. S., & Saneifard, R. (2021). A chaotic and hybrid gray wolf-whale algorithm for solving continuous optimization problems. *Progress in Artificial Intelligence*, 10(3), 349-374.
 25. Karthikeyan, M., Manimegalai, D., & RajaGopal, K. (2024). Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection. *Scientific Reports*, 14(1), 231.
 26. Hamad, R. K., & Rashid, T. GOOSE Algorithm: A New Bio-Inspired Metaheuristic Algorithm for Solving Optimization Problems Based on Goose Behavior. Available at SSRN 4581737.
 27. Hamad, R. K., & Rashid, T. A. (2023). GOOSE Algorithm: A Powerful Optimization Tool for Real-World Engineering Challenges and Beyond. *arXiv preprint arXiv:2307.10420*.
 28. Hamad, R. K., & Rashid, T. A. (2024). GOOSE algorithm: A powerful optimization tool for real-world engineering challenges and beyond. *Evolving Systems*, 15(4), 1249-1274.
 29. Khosrowshahi, H. N., Aghdasi, H. S., & Salehpour, P. (2025). A refined Greylag Goose optimization method for effective IoT service allocation in edge computing systems. *Scientific Reports*, 15(1), 15729.
 30. Alhussan, A. A., & Towfek, S. K. (2024). 5G Resource Allocation Using Feature Selection and Greylag Goose Optimization Algorithm. *Computers, Materials & Continua*, 80(1).
 31. Mostafa, R. R., Khedr, A. M., AL Aghbari, Z., Afyouni, I., Kamel, I., & Ahmed, N. (2025). A multi-strategy improved electric eel foraging optimization algorithm: continuous and binary variants for solving optimization problems. *International Journal of Machine Learning and Cybernetics*, 1-46.
 32. Mashru, N., Tejani, G. G., & Patel, P. (2025). Reliability-based multi-objective optimization of trusses with greylag goose algorithm. *Evolutionary Intelligence*, 18(1), 25.
 33. Hamad, R. K., & Rashid, T. GOOSE Algorithm: A New Bio-Inspired Metaheuristic Algorithm for Solving Optimization Problems Based on Goose Behavior. Available at SSRN 4581737.
 34. Hamad, R. K., & Rashid, T. A. (2024). GOOSE algorithm: A powerful optimization tool for real-world engineering challenges and beyond. *Evolving Systems*, 15(4), 1249-1274.
 35. Vadakkethil, S. E., Polimetla, K., Velpula, S., Pareek, P. K., & Sontakke, D. (2024). Improved Whale Optimization Algorithm and Optimized Long Short-Term Memory for DDoS Cyber Security Threat. In *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)* (pp. 01-05). IEEE.
 36. Dembele, A., Mwangi, E., Ronoh, K. K., & Ataro, E. O. (2024). A Novel Approach for Detection of DDoS Attacks in Software-Defined Networks Based on Grey Wolf Optimizer and Support Vector Machine. *SSRG International Journal of Electrical and Electronics Engineering*, 11(3), 86-96.
 37. Ramya, M., Sudhakaran, P., & Santhanakrishnan, C. (2025). Detecting DDoS Attacks for Enhancing Healthcare Data Security with Deep Neural Network (CNN-LSTM) and Particle Swarm Optimization. In *Utilizing AI of Medical Things for Healthcare Security and Sustainability* (pp. 401-424). IGI Global Scientific Publishing.
 38. Najar, A. A. (2024). A robust DDoS intrusion detection system using convolutional neural network. *Computers and Electrical Engineering*, 117, 109277.
 39. Mamatha, P., Balaji, S., & Anuraghav, S. S. (2025). Development of Hybrid Intrusion Detection System Leveraging Ensemble Stacked Feature Selectors and Learning Classifiers to Mitigate the DoS Attacks. *International Journal of Computational Intelligence Systems*, 18(1), 20.
 40. Bakro, M., Kumar, R.R., Husain, M., Ashraf, Z., Ali, A., Yaqoob, S.I., Ahmed, M.N. and Parveen, N., 2024. Building a Cloud-IDS by Hybrid Bio-Inspired Feature Selection Algorithms Along with Random Forest Model. IEEE Access
 41. Alharbi, A. H., Khafaga, D. S., El-kenawy, E. S. M., Eid, M. M., Ibrahim, A., Abualigah, L., ... & Abdelhamid, A. A. (2024). Optimizing electric vehicle paths to charging stations using parallel greylag goose algorithm and Restricted Boltzmann Machines. *Frontiers in Energy Research*, 12, 1401330.
 42. Beaumont, F., Murer, S., Bogard, F., & Polidori, G. (2024). The aerodynamic mechanisms of the formation flight of migratory birds: A narrative review. *Applied Sciences*, 14(13), 5402.