JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-21, No.-04, April (2026) pp 112-134 ISSN (Print) 0973-8975

EFFICIENT STATIC DISTRIBUTION AWARE TWO CLUSTER INTRUSION DETECTION SYSTEM FOR BINARY CLASSIFICATION USING DBF CLUSTERING AND PSO FEATURE SELECTION WITH MACHINE LEARNING MODELS

Hasan Abdulrazzaq Jawad [1], Shurooq M Abdulkhudhur [2], Rand A. Atta[3] Zahraa Ibrahim Abed[4]

[1] Department of Cyber Security, Imam AlKadhim University College (IKC), Baghdad, Iraq.

[2] College of Computer Science and Information Technology, University of Al-Qadisiyah, Al Diwaniyah, 58001, Iraq.

[3] College of Basic Education, Tikrit University, Sharqat, Iraq.

[4] College of Health and Medical Technologies, Department of Radiology.

Email: [1]hasan.abdulrazzaq@iku.edu.iq, [2]shurooq.alrubaye@qu.edu.iq, [3]Rand.a.ata@tu.edu.iq, [4]zahraa.ibraheem@alzahraa.edu.iq

Corresponding author's Email: **Hasan Abdulrazzaq Jawad**

https://doi.org/10.26782/jmcms.2026.04.00006

(Received: January 27, 2026; Revised: March 26, 2026; Accepted : April 08, 2026)

Abstract

Network protection relies on machine learning-based systems that detect intrusions. The detection systems lose their effectiveness because they use multiple duplicate features, and their performance depends on the specific network traffic patterns and system operational requirements, which prevent real-time functioning. The research presents a PSO-DBF intrusion detection framework, which begins with Distributional Boosting Forest (DBF) as its first step to create two groups (C1 and C2) that display similar probabilistic characteristics through network traffic clustering. The research team uses Particle Swarm Optimization (PSO) to process each cluster when they complete their clustering process because the method helps them find the most valuable network attributes, which decrease feature duplication while enhancing the ability to distinguish different features. K-Nearest Neighbors (KNN) provides the best performance when conventional machine learning classifiers use optimized feature subsets for intrusion detection. The proposed framework demonstrated its efficiency through experiments that utilized recognized IDS datasets. PSO removed almost 50% of the initial features while keeping 18 features from NSL-KDD and 21 features from UNSW-NB15, achieving reduction rates of approximately 56 percent and 57 percent. The proposed PSO–DBF with KNN framework achieved 99.36% accuracy on NSL-KDD and 99.89% accuracy on UNSW-NB15, exceeding the

performance of Support Vector Machine (SVM), Naive Bayes (NB), Quadratic Discriminant Analysis (QDA), deep neural models, and recent hybrid metaheuristic-based IDS frameworks. The main improvement of the proposed method comes from its ability to reduce detection times, which drop from 0.44 milliseconds to 0.29 milliseconds. The DBF-PSO framework achieves its optimal performance for intrusion detection in enterprise cloud and edge-network security environments because of its detection accuracy and energy efficiency.

Keywords: Network Security Intrusion Detection System, Particle Swarm Optimization, Distributional Boosting Forest, Machine Learning, Cyberattack Detection, Real-Time Threat Monitoring.

I. Introduction

The modern digital world needs intrusion detection systems because its growing dependence on linked digital systems creates security weaknesses that allow cybercriminals to launch attacks and gain unauthorized access, thereby compromising data protection and service reliability [I, XIII]. IDS solutions analyse traffic behaviour to differentiate between legitimate communication and malicious events; however, these days, greater challenges are associated with identifying abnormal patterns as network environments continue to expand and cyber threats continue to evolve [XX, XXXII].

With machine learning, which is capable of modelling complex patterns and adapting to new forms of intrusions in its capacity as an emerging approach in enhancing intrusion detection system performance, these classical models, such as SVM [II], NB [XVIII], QDA [IX], have further proven their lightweight and scalability. Reduced computational demand [XIX] does not imply reduced performance due to the fact that many of these classical ML models tend to perform poorly when subjected to high-dimensional feature spaces or when noisy input signals are presented. When dynamic behaviours occur in the networks, these factors lead to increased inconsistencies in detection accuracy and increased false alarms [XXI], [III].

With a view to tackling the presented issues, this study proposes an efficient adaptive intrusion detection framework using DBF–PSO. The proposed framework uses DBF to create probabilistic traffic distribution models, which include uncertainty features for better distribution-based clustering than K-means, which uses distance measurements to group network traffic. In the proposed design, DBF is first employed to cluster network traffic into two static distribution-aware clusters, enabling a clearer separation between normal and malicious behaviours in a binary classification setting. Subsequently, PSO optimally identifies and retains the most relevant traffic features within each cluster by eliminating redundancy, thereby enhancing the discriminative capability and efficiency of the downstream machine learning classifiers. The following are the major contributions of this paper:

- An adaptive hybrid IDS architecture is introduced that DBF partitions network traffic into two distribution-aware clusters (C1 and C2), enabling structured separation of normal and malicious behaviours. PSO then selects the most

informative features within each cluster, after which machine learning classifiers perform the final intrusion classification.

- The feature set is reduced from 41 to 18 features (56% reduction) for NSL-KDD and from 49 to 21 features (57% reduction) for UNSW-NB15, effectively eliminating redundant and noisy attributes.
- DBF–PSO–KNN approach, which was developed, showed 99.89% accuracy on UNSW-NB15 while achieving 99.36% accuracy on NSL-KDD assessments to surpass QDA, SVM, NB, and modern hybrid IDS systems. The detection latency reduction from 0.44 ms to 0.29 ms shows improved results, which enable real-time monitoring of network security.

II. Related Works

The current state of IDS research employs feature reduction techniques together with particular machine learning methods to achieve better detection results while decreasing the operational demands of processing power. The existing literature currently shows a strong consensus that supports the use of metaheuristic optimization methods to choose small yet distinctive feature sets that should be used for classification tasks. The research study conducted by Umar et al. [XXVIII] analyzed how two different techniques, min-max normalization and wrapper-based feature selection, affected the performance of various classifiers. The tree-based classifiers experienced the most advantages from feature reduction, whereas neural networks showed improvements through normalization. The study did not include any testing to measure how well these configurations would perform in actual real-time IDS operations.

In this regard, Hammood et al. [XIV] implemented XGBoost on the UNSW-NB15 dataset as a classifier and incorporated PSO to find the best subset of features. The system showed almost 99.51% accuracy with low false alarm rates. However, they did not consider computational overhead during training and inference, which is important from a large deployment perspective. Also, Zhixin Xia et al. [XXX] put forth a hybrid deep model combining BiGRU architecture and ResNet, where parameters were tuned through combined PSO-GA search. The technique was successful in achieving an accuracy of 98.46% on UNSW-NB15, while no deliberations were made in regard to such factors as computational delay and load incurred on the system.

The method proposed for the NSL-KDD dataset by Raghunath et al. [XXVII] aimed to improve classification precision, and the SVM was chosen as the core of the model, while PSO was used to automatically find relevant features. To simplify the input space, they performed normalization as a PCA pre-process. Their model achieved 98.5% accuracy, outperforming a number of baseline learning algorithms. Unfortunately, however, the verification was limited to a single dataset, and as such, its generalization capability remains questionable.

Deep learning-based IDS models have also started to couple optimization techniques. Yılmaz [XXXI] experimented with pre-trained neural networks, where hyperparameter tuning was done using PSO. For NSL-KDD and UNSW-NB15, an accuracy of 90.42% and 82.24% was obtained, respectively; however, the work did not consider inference complexity, thereby limiting its applicability for real-time intrusion detection scenarios. As stated in the study done by Kurdi et al. [XXIX], after first clustering

network traffic with K-means, application of Cuckoo search (CS) was made in order to extract the most vital features from each cluster. The model was tested on the NSL-KDD and UNSW-NB15 datasets and achieved 98.7% and 99.78% accuracy, respectively, with significantly fewer features. However, the authors did not present an analysis of execution cost, and thus the approach's real-time applicability remains unknown. Chakravarty et al. [XI] implemented LSTM networks to analyze the temporal behavior of attacks and optimized it using SSA, PSO, and JAYA. The highest performance was recorded on the NSL-KDD at 97.89% accuracy using the SSA-tuned model. However, the study did not entertain the size of the final feature subset or conduct the evaluation under real-time throughput conditions. In [XII], an evaluation of several ML models coupled with optimization algorithms, Gradient Boosting with Differential Evolution, RF with Flower Pollination, and DT with PSO and DE, was more established and evaluated to check the performance on both NSL-KDD and UNSW-NB15. Their investigation indicated that optimization is viable in achieving a significant wracking of F1-scores along with saving computation time. Nonetheless, the modern characteristics of these methods regarding scalability in high-speed networks weren't explored.

The research on intrusion detection has used feature selection methods together with clustering methods, yet most studies depend on traditional clustering methods, which include K-means. The K-means algorithm divides data into clusters by using distance measurements, which most commonly use Euclidean distance to create groups based on their geometric resemblance in their feature vectors. The method offers high computational efficiency, but it fails to accurately represent the statistical distribution and uncertainty aspects that define network traffic patterns during complex dynamic conditions. The proposed framework uses DBF as its distribution-aware clustering mechanism for network traffic clustering. Distance-based clustering uses distances for clustering, while DBF uses probabilistic models together with boosting distribution estimation to create direct data processing of traffic patterns and their accompanying uncertainty. The clustering process generates traffic groups that match specific distribution patterns, thus enhancing the performance of PSO-based feature selection and boosting the classification accuracy of the final classification models.

III. Suggested Methodology

The proposed IDS adopts a hierarchical and distribution-aware pipeline in which network traffic from the NSL-KDD and UNSW-NB15 datasets is first normalized using Unit Vector Normalization. A DBF is then used as a clustering mechanism to divide the traffic into two homogeneous clusters (C1 and C2) representing normal-dominant and suspicious-dominant behaviours. PSO is applied separately within each cluster to select the most informative features, after which QDA, SVM, NB, and KNN classifiers are employed for intrusion detection. The model is evaluated using standard performance metrics. The workflow of the proposed approach is shown in the Fig. 1.

III.i. Dataset Description

The choice of datasets that represent different elements of network traffic patterns proves to be essential for establishing the effectiveness of intrusion detection systems.

The research team selected NSL-KDD [XXII] and UNSW-NB15 [XXV] because these two datasets reflect different time periods of network traffic; their selection defines modern attack research and traditional attack research. The UNSW-NB15 dataset contains 257673 traffic instances, which include 175341 training samples and 82332 testing samples. The dataset includes 49 attributes that describe all records of normal traffic, together with 10 different types of attack traffic. The dataset provides current intrusion detection methods because it shows how detection systems handle emerging threat patterns. The NSL-KDD dataset contains 148517 records, which include 22544 testing records and 125973 training records. The dataset functions as a better version of KDD'99 because it eliminates duplicate data and creates an equal distribution of classes, which enables machine learning models to learn through fair training practices. The combination of both datasets enables researchers to achieve dual assessment benefits because UNSW-NB15 tests current advanced intrusions while NSL-KDD provides a fair benchmark to evaluate classification accuracy. The use of both datasets strengthens the proposed IDS system because it improves its ability to detect and identify intrusions. Table 1 presents a summary of all dataset characteristics.

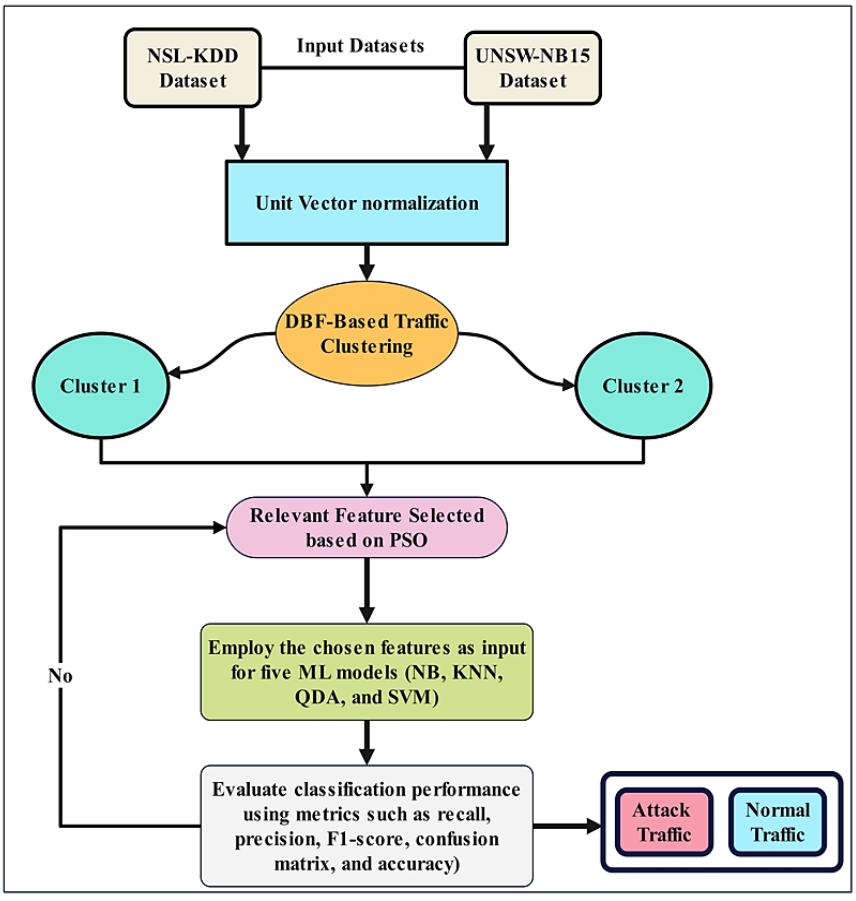


Fig. 1. Workflow of the distribution-aware IDS utilizing DBF clustering and PSO-driven feature selection with ML classifiers.

Table 1. IDS Datasets Details.

Dataset	Source and Year	Number of Samples	Number of Features	Attack Categories
NSL-KDD	Improved KDD'99 Release, 2009	148,517	41	5
UNSW-NB15	Cyber Range Lab, 2015	257,673	49	10

III.ii. Pre-processing Network Traffic

The different scales and values present in raw traffic data1 lead to biased learning behavior1, which affects the process of categorization. The learning process gets dominated by attributes that have big values, while attributes with small values get completely disregarded according to this research study [IV, V]. The IDS model training process requires all features to have their representation standardized before any training begins. The research study uses Unit Vector Normalization as its method for feature normalization. The transformation process changes each feature vector into a unit-length version, which allows every sample to contribute equally, regardless of its original size. Let $X = [x_1, x_2, x_3, \ldots\ldots, x_n]$ represent the original feature vector for a network record. The normalized vector X' is computed as:

$$\boldsymbol{X'} = \frac{X}{||X||} = \frac{X}{\sqrt{\sum_{i=1}^{n} X_i^2}} \tag{1}$$

III.iii. DBF-Based Traffic Clustering

Intrusion detection faces challenges because normal and malicious network traffic patterns show overlapping similarities, which become especially difficult to distinguish in dynamic network systems that operate at large scales. Conventional distance-based clustering methods fail to properly capture the uncertainty that exists in network traffic patterns. The DBF framework functions as a probabilistic clustering system that enables traffic classification based on distributional similarities while it models class probabilities and prediction uncertainty [VIII].

Let (X_i, y_i) represent a network traffic sample, where X_i denotes the feature vector and $y_i \in \{0, 1\}$ indicates normal or attack behavior. DBF estimates the probability of intrusion as:

$$\mathbf{P(y_i = 1|X_i)} = \Phi(f(X_i)) \tag{2}$$

Where: Φ is the cumulative distribution function of the standard normal distribution and $f(X_i)$ represents the aggregated contribution of the boosted trees. The objective of training minimizes the following loss function:

$$\mathbf{L} = \sum_{i=1}^{n} [y_i \log \Phi(f(X_i) + (1 - y_i) \log(1 - \Phi(f(X_i)))] \tag{3}$$

DBF models a variance term $\alpha^2(X_i)$ because it needs to predict outcomes while using its probability assessment capabilities. The model predicts outcomes through its probability assessment and variance term function:

$$\hat{\mathbf{y}}_{\mathbf{i}} \sim (\mathrm{f}(\mathrm{X_i}), \alpha^2(\mathrm{X_i})) \quad (4)$$

Network traffic is grouped into two clusters based on computed confidence intervals or prediction limits:

- Cluster 1: High-confidence normal-dominant traffic
- Cluster 2: High-uncertainty or attack-dominant traffic

The main purpose of this clustering step is to bring the classification of instances to an end, but only traffic loads to enhance the effect of the subsequent select-feature and single-step classification components.

Output Classification Decision: For a new traffic input X^*, the probability of belonging to the intrusion class is determined as:

$$\mathbf{p} = \mathrm{P}(\mathrm{y_i} = 1|\mathrm{X}^*) = \Phi(\mathrm{f}(\mathrm{X}^*)) \quad (5)$$

The final decision is controlled through an adjustable threshold τ, defined as:

$$\mathbf{Class}(\mathbf{X}^*) = \begin{cases} \text{Normal} & \text{if } \mathrm{p} < \tau \\ \text{Attack} & \text{if } \mathrm{p} \geq \tau \end{cases} \quad (6)$$

The combined prediction result is produced through the process of blending all the individual predictions that have been generated by the boosted trees.

$$\mathrm{f}(\mathrm{X}^*) = \frac{1}{\mathrm{T}} \sum_{\mathrm{t}=1}^{\mathrm{T}} \mathrm{h_t}(\mathrm{X}^*) \quad (7)$$

Where $\mathrm{h_t}$ denotes the contribution of the $\mathrm{t^{th}}$ learner. The DBF system enables accurate classification results while delivering prediction confidence levels that enhance intrusion detection performance under conditions of network uncertainty. The algorithm 1 shows the distribution-aware IDS framework that proposed in this study.

III.iv. Relevant Feature Choices Based on PSO

High-dimensional network traffic data typically contains redundant and irrelevant attributes that negatively impair the detection speed and classification accuracy [XV]. Selecting the most informative features will thus reduce computational overhead while maximizing the performance of an IDS. PSO is applied in this work to give an optimal selection of features from this original dataset. The PSO, with its collective motion of bird flocks, is a population-based stochastic optimization technique [XVI]. In PSO, multiple particles represent candidate solutions and move throughout the solution space. Every particle stores its own best location $\boldsymbol{P_{best}}$ and the swarm-wide best location G_{best}. The position and velocity of each particle are updated during the search process as follows:

$$\boldsymbol{v_i^{(t+1)}} = Q\, v_i^{(t)} + \alpha_1\, m_1(\boldsymbol{P_{best}} - S_i^{(t)}) + \alpha_2\, m_2\, (\boldsymbol{G_{best}} - S_i^{(t)}) \quad (8)$$

$$\boldsymbol{S_i^{(t+1)}} = \boldsymbol{v_i^{(t+1)}} + \boldsymbol{S_i^{(t)}} \quad (9)$$

Where: $S_i^{(t)}$ is current position (feature subset) of the particle, $v_i^{(t)}$is velocity of the particle, Q is Inertia weight controlling exploration/ exploitation, $\alpha_1$and α_2 serve as the personal experience and social cooperation coefficients, m_1 and m_2 are random

numbers in the range [0, 1]. A fitness function is a measure of merit for candidate feature subsets [XXIII]. In the framework proposed in this work, the fitness function is centred on the classification accuracy in that it retains only the features that contribute positively to the intrusion recognition. In this manner, particles gradually converge on the subset that maximizes detection performance while minimizing inclusion of irrelevant features. The use of PSO for feature selection would allow for a more compact representation of traffic patterns and a more manageable input space, in turn improving training and detection reliability. The fitness function is given by:

$$\boldsymbol{Fitness(S)} = \textit{Accuracy of ML models} \tag{10}$$

	Algorithm 1: Distribution-Aware IDS Framework
	Input: **D_Net:** Network traffic datasets (NSL-KDD and UNSW-NB15). **C_No:** Number of clusters (set to 2 for Normal-Dominant and Suspicious-Dominant). **P_Size:** Number of particles in the PSO population. **Iter:** Number of iterations for feature selection.
	Output: **Acc_Results:** A report containing performance metrics for QDA, SVM, NB, and KNN.
1	**Phase 1: Preprocessing** -For each record in **D_Net**: • Apply Unit Vector Normalization to scale values. -End for
2	**Phase 2: Distribution-Aware Clustering (DBF)** -For each normalized record (i = 1 to Total_Records): • Apply DBF to determine the data distribution. • Calculate similarity to cluster centers. • Assign record to Cluster_0 (**Normal-Dominant**) or Cluster_1 (**Suspicious-Dominant**). -End for -Save results into two separate data pools: **Pool_Normal** and **Pool_Suspicious**.
3	**Phase 3: Cluster-Specific Feature Selection (PSO)** -For each Cluster (j = 1 to **C_No**): 1- Initialize PSO particles with random feature subsets. 2-For each iteration (t = 1 to **Iter**): • Evaluate fitness based on classification error. • Update Global Best (**Gbest**) and Personal Best (**Pbest**) feature sets. 3- End for 4-Extract the **Best_Feature_Subset** for that specific cluster. -End for
4	**Phase 4: Classification and Evaluation** -Apply the selected features to the Classifiers: QDA, SVM, NB, and KNN. -Compute performance metrics (Accuracy, F1- Score, Detection Rate). -Return **Acc_Results**.

III.v. Evaluation Metrics for ML Models

With higher precision, the capability of differentiating normal and malicious traffic is highlighted for evaluating the performance of an IDS. The present study, therefore, considers four common performance measures, Recall, Precision, F1-score, and

Accuracy, to assess the efficacy of various classification strategies in distinguishing normal traffic from malicious traffic. Also serves diligently for identifying where the crop classification is efficient, especially where the classes are imbalanced [XXIII, X].

- Precision is a score that shows how well it detects harmful traffic from all samples it marks as harmful. The presence of fewer false alarms is shown through increased precision results [XXIV, XXVI].

$$\textbf{Precision} = \frac{TP}{TP + FP} \tag{11}$$

- Recall refers to the model's ability to find real criminal activities, which is measured through its recall performance. A high recall score indicates that fewer attacks go undetected [VI].

$$\textbf{Recall} = \frac{TP}{TP + FN} \tag{12}$$

- The F1-Score is defined as the harmonic mean of Precision and Recall, providing a useful balance when both false detection and missed detection have to be minimized [XVII].

$$\textbf{F1} - \textbf{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Recall} + \text{Precision}} \tag{13}$$

- Accuracy assesses correctly classified samples in relation to the total number of samples and represents a summary measure of the model's performance [VII].

$$\textbf{Accuracy} = \frac{TN + TP}{FN + FP + TN + TP} \tag{14}$$

IV. Experimental Outcomes and Critical Evaluation

The proposed intrusion detection framework was evaluated on the NSL-KDD and UNSW-NB15 datasets to examine the effectiveness of the hierarchical (DBF–PSO–ML) approach. The distribution-aware clustering system DBF separates network traffic into two uniform groups, which PSO uses to determine the most valuable features from each group. The feature subsets that were optimized through the process serve as the training material for the classifiers QDA, SVM, NB, and KNN. The evaluation of classifiers requires all systems to use the same PSO-selected features, which are present in each cluster. The study found that DBF-based clustering combined with cluster-wise PSO enhances classification achievement for all classifiers through both datasets. The clustering process decreases traffic variation, while feature selection helps remove unnecessary elements, which results in better detection performance and reduced false alarm rates. The DBF-PSO hyperparameter values and classifier settings appear in Table 2 to enable researchers to recreate the study results. PSO-based feature selection reduces input features, which results in decreased computational power requirements and memory needs for classification tasks, thus enabling the framework to function in cloud and edge environments with limited resources.

Table 2: Hyperparameter settings used in the experiments.

Component	Parameter
DBF (Clustering)	Number of Trees= 200, Learning Rate=0.05, Loss Function= Probit-based Boosting Loss
PSO	Swarm Size = 30, Maximum Iterations = 100, Inertia Weight= 0.72, Cognitive & Social Coefficients= α_1 =α_2= 1.4
QDA	Covariance Structure= Full Covariance, Regularization= 0.0001
SVM	Kernel Type= Radial Basis Function (RBF), Regularization Constant= 1, Kernel Scale= 0.01
NB	Distribution Type= Gaussian NB, Variance Smoothing= 1×10−3
KNN	Number of neighbors = 5, Distance metric = Euclidean

IV.i. ML Models Results on the NSL-KDD With and Without BFD-PSO

The machine learning models QDA, SVM, NB, and KNN showed classification performance results on the NSL-KDD dataset in Table 3, which showed results before and after researchers applied PSO-based feature selection to DBF-generated clusters. DBF operates in this system to create two identical network traffic clusters while PSO conducts feature optimization for each cluster before the classification process. The classifiers function with their complete feature set for every cluster when PSO is not implemented. The KNN method achieved the highest accuracy at 95.7% in this experiment, while QDA and NB, and SVM achieved accuracy rates of 94.32% and 91.67% and 91.4%, respectively. The performance decrease of some classifiers results from redundant features, which provide less information because they disrupt decision boundaries and lead to more errors in classification.

Table 3: Performance comparison on the NSL-KDD of the proposed BDF-PSO with four ML.

Without (BDF-PSO)				
Models	**Accuracy**	**F1-Score**	**Recall**	**Precision**
QDA	94.32	93.15	92.6	93.7
SVM	91.4	90.34	89.6	91.1
NB	91.67	90.5	91.1	89.9
KNN	**95.7**	**94.45**	**95.1**	**93.8**
With (BDF-PSO)				
Models	**Accuracy**	**F1-Score**	**Recall**	**Precision**
QDA	98.5	98.5	98.55	98.46
SVM	98.44	98.45	98.32	98.58
NB	98.41	98.42	98.32	98.52
KNN	**99.36**	**99.36**	**99.38**	**99.35**

The application of PSO leads to improved performance across all classifiers because it eliminates unnecessary and duplicated features. The highest accuracy is reached by KNN (99.36%), followed by QDA (98.50%), SVM (98.44%), and NB (98.41%). The classification behavior of the best-performing model (KNN) on the NSL-KDD test set is illustrated in Fig. 2. The confusion matrix shows that 3358 normal samples and 3337 attack samples are correctly classified, with only 22 normal instances and 21 attack instances misclassified. The proposed DBF-based clustering with PSO feature selection demonstrates reliable and consistent classification results according to the confusion matrix because the method enables the detection of attacks that closely imitate normal traffic patterns.

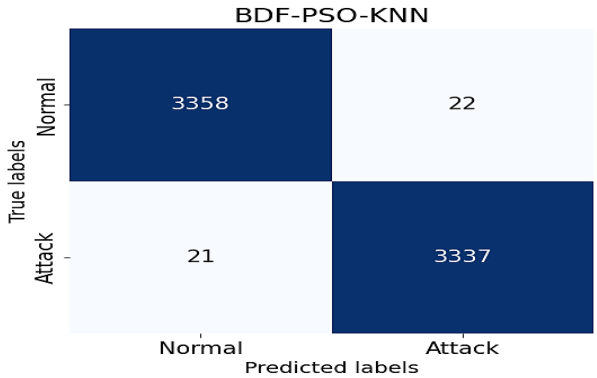


Fig. 2. Confusion Matrix of the DBF-PSO-KNN Approach on NSL-KDD.

IV.ii. ML Models Result on the UNSW-NB15 with and without BDF-PSO

Evaluation of the proposed IDS framework through contemporary sophisticated attack patterns received assessment from the second benchmark, which used the UNSW-NB15 dataset. The table displays the classification results for QDA, SVM, NB, and KNN before and after the implementation of PSO-based feature selection, which operates on DBF-generated clusters. The classifiers function on their complete feature set, which follows DBF-based clustering when PSO remains inactive. The KNN algorithm attained maximum accuracy with a result of 96.55% which the QDA algorithm followed with 95.94% accuracy, the SVM algorithm with 95.23% accuracy, and the NB algorithm with 94.15% accuracy. DBF clustering organizes traffic flows through its clustering process. However, the existence of unnecessary features together with less valuable features creates obstacles to successful classification, especially during complex attacks that share similarities with other attacks. The implementation of PSO across every cluster managed to deliver substantial enhancements to all classifiers. KNN achieved the best accuracy rate of 99.89%, while QDA followed with 99.54% accuracy, and SVM came third with 99.24% accuracy, and NB achieved 99.16% accuracy. The detection system maintained its balanced detection performance through the close relationship between Precision and Recall and F1-score values, while cluster-wise feature optimization created better separation between normal traffic and malicious traffic.

Table 4: Performance comparison on the UNSW-NB15 of the proposed BDF-PSO with four ML.

Without (BDF-PSO)				
Models	**Accuracy**	**F1-Score**	**Recall**	**Precision**
QDA	95.94	95.76	95.41	96.12
SVM	95.23	95.18	94.81	95.55
NB	94.15	94.10	94.52	93.68
KNN	**96.55**	**96.57**	**96.82**	**96.33**
With (BDF-PSO)				
Models	**Accuracy**	**F1-Score**	**Recall**	**Precision**
QDA	99.54	99.35	99.29	99.4
SVM	99.24	98.93	98.72	99.14
NB	99.16	98.83	98.72	98.94
KNN	**99.89**	**99.84**	**99.84**	**99.85**

The combination of DBF-based clustering with PSO feature selection produces a compact and distinct model of network traffic, which enables standard machine learning classifiers to achieve high detection rates throughout various types of intrusions. The confusion matrix for the best-performing classifier (KNN) on the UNSW-NB15 test set is shown in Fig. 3. The model correctly classified 27,236 normal samples and 49,305 attack samples, with only 41 normal instances misclassified as attacks and 45 attack instances misclassified as normal. The system demonstrates its ability to discriminate between samples and identify correct samples because of its extremely low erroneous identification rate.

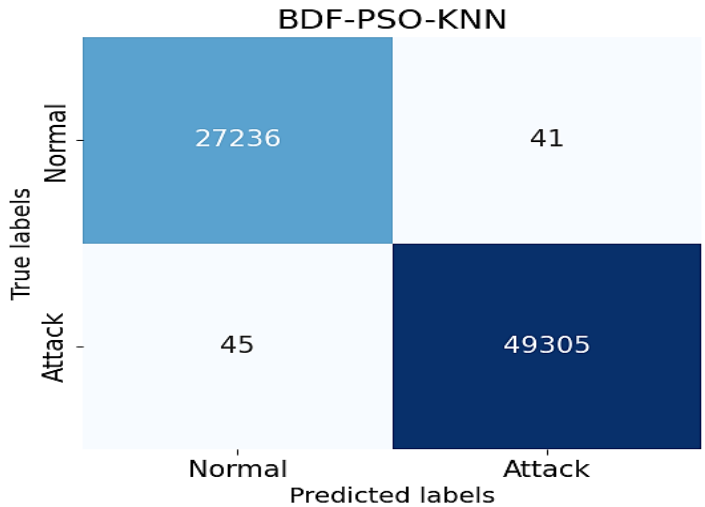


Fig. 3. Confusion Matrix of the BDF-PSO-KNN Approach on UNSW-NB15.

Table 5: Average Latency per Traffic Sample (ms) with and without BDF-PSO.

Model	NSL-KDD		UNSW-NB15	
	(41 features)	(BDF-PSO, 16 features)	(49 features)	(BDF-PSO, 19 features)
QDA	0.48	0.34	0.46	0.31
SVM	0.57	0.41	0.53	0.38
NB	0.25	0.18	0.24	0.17
KNN	0.44	0.29	0.42	0.27

IV.iii. Runtime Evaluation and Complexity Assessment

The operation of intrusion detection systems depends on their ability to detect intrusions within real-time network environments that handle high data traffic. Table 5 and Fig. 4 show detection latency results for QDA, SVM, NB, and KNN, which were tested on NSL-KDD and UNSW-NB15 datasets before and after PSO-based feature selection was applied to DBF-generated clusters. The classifiers function with complete feature sets, which include 41 features for NSL-KDD and 49 features for UNSW-NB15, after the DBF-based clustering process. The NB classifier achieves the shortest inference duration in this configuration because its basic probabilistic model requires less processing time, while SVM and QDA take longer because their decision-making processes require more complicated algorithms. KNN has an average inference duration because its classification process needs to calculate distances to determine the output result. The PSO function decreased the feature space from 41 elements to 16 elements for NSL-KDD and from 41 elements to 19 elements for UNSW-NB15. The reduction of feature space results in classifiers needing less time to detect targets. KNN decreased its runtime on NSL-KDD from 0.44 ms to 0.29 ms and on UNSW-NB15 from 0.42 ms to 0.27 ms while maintaining optimal detection performance combined with quick processing times. The same improvements were achieved by QDA, SVM, and NB. The results show that PSO feature selection at the cluster level decreases computational demands because it removes unneeded features, while the DBF clustering process only requires a small amount of time for its initial data processing. The DBF–PSO–ML framework provides quick and precise intrusion detection for use in high-throughput network environments which operate in real time.

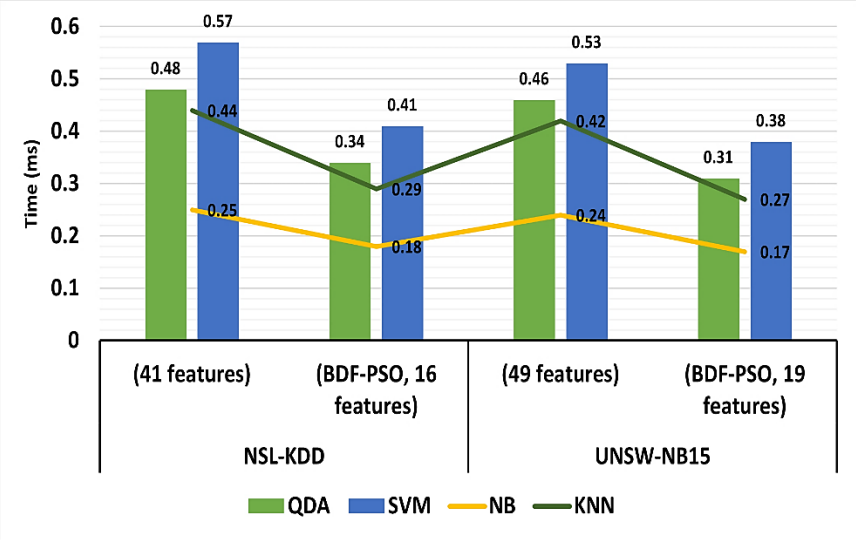


Fig. 4. Average per-sample inference time using distribution-aware IDS approach (DBF-PSO-ML).

PSO algorithm identifies the most informative traffic attributes while eliminating redundant ones. The selected features for both datasets are summarized in Table 6, which improves the interpretability of the proposed intrusion detection framework.

Table 6. Selected features obtained by PSO for both two IDS datasets.

Dataset	Selected Features using PSO
NSL-KDD	protocol_type, src_bytes, duration, dst_bytes, service, flag, srv_count, count, logged_in, same_srv_rate, dst_host_srv_count, dst_host_count, dst_host_diff_srv_rate, dst_host_same_srv_rate, dst_host_srv_serror_rate, dst_host_serror_rate
UNSW-NB15	proto, dur, service, state, dbytes, sbytes, dload, rate, sload, sttl, dttl, smean, dmean, ct_state_ttl, ct_dst_src_ltm, ct_srv_src, ct_src_dport_ltm, ct_dst_ltm, is_sm_ips_ports

IV.iv. Classifier Performance Analysis Using Various Metaheuristic Algorithms

The study evaluates how different metaheuristic feature selection algorithms, which include PSO, GWO, WOA, and GA, impact the DBF-based clustering framework when applied to the NSL-KDD and UNSW-NB15 datasets. The DBF method is applied first to create two homogenous clusters of network traffic, which serve as the starting point for all experiments. The clusters use each metaheuristic algorithm to pick the best feature subsets, which lead to classification tests that employ identical machine learning models for unbiased assessment. The quantitative results are presented in Table 7. The PSO method reaches its peak performance on the NSL-KDD dataset by achieving an F1-score of 99.36% and an accuracy of 99.36%. GWO achieves competitive performance through its 99.28% accuracy and 99.25% F1-score results. The two methods achieved lower performance results, with WOA achieving a 99.12% accuracy, and GA achieving a 98.87% accuracy. PSO demonstrates superior

performance in identifying compact and discriminative feature subsets because it shows numerical advantages that, although minor, consistently make it better than all other options. The UNSW-NB15 dataset reveals an identical pattern since it displays advanced attack techniques that are more difficult to detect. PSO again provides the best performance, achieving an accuracy of 99.89% and an F1-score of 99.84%. GWO attains 99.76% accuracy and 99.71% F1-score, while WOA and GA reach 99.62% and 99.41% accuracy, respectively. The consistent ranking of the metaheuristic algorithms across both datasets confirms the robustness and generalization capability of the proposed framework, as shown in Table 7 and Fig 5.

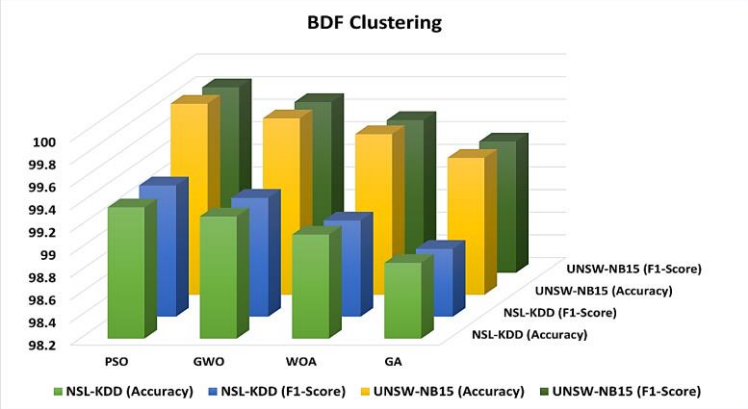


Fig. 5. Accuracy and F1-score comparison of PSO, GWO, WOA, and GA after DBF-based traffic clustering.

Table 7: DBF performance with different metaheuristics on NSL-KDD and UNSW-NB15.

Metaheuristic (with DBF clustering)	NSL-KDD (Accuracy)	NSL-KDD (F1-Score)	UNSW-NB15 (Accuracy)	UNSW-NB15 (F1-Score)
PSO	99.36	99.36	99.89	99.84
GWO	99.28	99.25	99.76	99.71
WOA	99.12	99.05	99.62	99.55
GA	98.87	98.80	99.41	99.36

IV.v. Significance Testing and Performance Validation

Paired t-tests were used to perform statistical significance testing, which showed that DBF-based clustering method improved performance before PSO feature selection. The researchers tested the accuracy of each classifier, which used and did not use DBF clustering, at a 95% confidence level. The QDA accuracy on the NSL-KDD dataset increased from 97.30% to 98.90% (+1.60%, $p = 0.003$), the SVM accuracy increased from 97.40% to 98.84% (+1.44%, $p = 0.002$), and the NB accuracy increased from 97.60% to 98.71% (+1.11%, $p = 0.001$). The QDA accuracy on the UNSW-NB15 dataset increased from 99.54% to 99.89% (+0.35%, $p = 0.004$), the SVM accuracy

increased from 99.24% to 99.89% (+0.65%, p = 0.003), and the NB accuracy increased from 99.16% to 99.89% (+0.73%, p = 0.003). The statistical results demonstrate that all p-values remain below 0.05, which proves that DBF-based clustering improvements reach statistically significant levels. The DBF-assisted framework demonstrated lower performance variation through its repeated tests, which showed that it provided better detection stability and reliability. The results presented in Table 8 demonstrate that DBF-based clustering combined with PSO optimization produces statistically significant performance improvements across both testing datasets. The statistical analysis shows that DBF-based clustering improves PSO-optimized intrusion detection performance across all tested classifiers and datasets, proving its reliability and practical value.

Table 8: Statistical Significance of DBF-Based Clustering on PSO-Optimized Machine Learning Models (Paired t-Test).

Models	Dataset	Accuracy without DBF	Accuracy with DBF	Mean Difference	p-Value
QDA	NSL-KDD	97.30	98.90	+1.60	0.003
KNN	NSL-KDD	98.05	99.36	+1.31	0.05
SVM	NSL-KDD	97.40	98.84	+1.44	0.002
NB	NSL-KDD	97.60	98.71	+1.11	0.001
QDA	UNSW-NB15	99.54	99.89	+0.35	0.004
SVM	UNSW-NB15	99.24	99.89	+0.65	0.003
KNN	UNSW-NB15	99.55	99.89	+0.34	0.05
NB	UNSW-NB15	99.16	99.89	+0.73	0.003

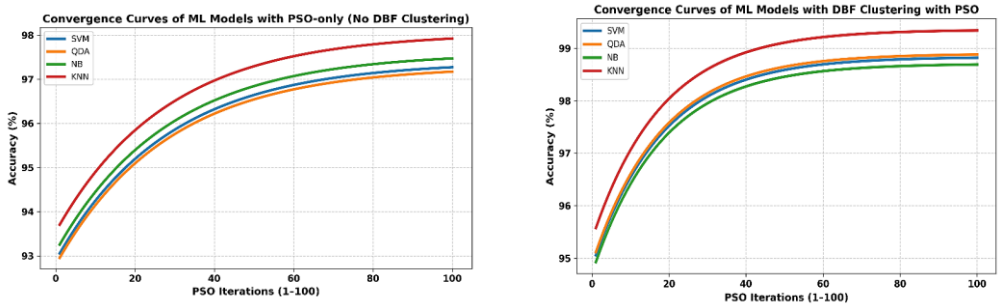
IV.vi. Effect of DBF-Based Clustering on PSO-Optimized Machine Learning Performance

The study examines how traditional machine-learning classifiers, which include QDA, SVM, NB, and KNN, perform when researchers use DBF-based traffic clustering to create new features before proceeding with PSO-based feature selection. The researchers tested 100 PSO iterations on all classifiers to avoid biasing results while they evaluated classifier performance through two testing methods, which included PSO feature selection without DBF clustering and the proposed framework that first conducts DBF clustering before PSO optimization. The experiment results demonstrate that DBF clustering provides additional benefits to PSO-based feature selection, which remains impossible to quantify for all classifiers. The final classification accuracies achieved through QDA SVM, NB, and KNN resulted in 97.30% 97.40% 97.60% and 98.05%, respectively. The accuracies achieved through DBF-based clustering before PSO optimization reached stable values of 98.90% for QDA, 98.84% for SVM, 98.71% for NB, and 99.36% for KNN. The accuracies established above achieved percentage increases of 1.60% for QDA, 1.44% for SVM, 1.11% for NB, and 1.31% for KNN, respectively. Our analysis of Table 9 demonstrates that DBF clustering has significantly improved the efficiency of PSO optimization.

Table 9. Impact of DBF-Based Clustering on PSO-Optimized Classification Accuracy.

Classifier	PSO-only Accuracy (%)	DBF+PSO Accuracy (%)	Accuracy Gain (%)
QDA	97.30	98.90	+1.60
SVM	97.40	98.84	+1.44
NB	97.60	98.71	+1.11
KNN	98.05	99.36	+1.31

The convergence analysis shown in Fig. 6 demonstrates that DBF-based clustering produces more consistent PSO optimization results, which enable classifiers to reach their best performance after fewer iterations with less performance fluctuations because DBF clustering decreases both intra-class variance and feature-space uncertainty, which leads to better results in feature selection. QDA and KNN both benefit from performance gains because distributionally homogeneous clustering enables them to extract more coherent feature representations. The results in Fig. 6 show that DBF-based clustering improves PSO feature selection accuracy while providing faster convergence and stronger intrusion detection capabilities for all tested classifiers.



(A) **(B)**

Figure 6. Convergence behaviour of classification accuracy over 100 PSO iterations for ML Models under: (A) PSO-only and (B) DBF with PSO configurations.

IV.vii. Analysis of Static Two-Cluster Structure and Selected Features

The DBF system creates two permanent clusters, C1 and C2, which show distribution patterns for network traffic before they start feature selection and classification. The NSL-KDD dataset contains 6738 samples, which are divided into two clusters with 3369 samples assigned to each cluster. The UNSW-NB15 dataset has 76627 samples, which are split into two equal parts of 38313 C1 samples and 38314 C2 samples. After the clustering process, each cluster uses PSO to find compact feature subsets that can distinguish between different groups of features. The NSL-KDD dataset shows that PSO reduced the feature space to 16 features per cluster, while the UNSW-NB15

dataset shows a reduction to 19 features per cluster, according to Fig 7. The cluster-specific feature selection method removes unnecessary features while keeping the most important features for each traffic pattern. The static two-cluster structure helps PSO-based feature optimization work better because it decreases intra-cluster variability and creates an easier learning environment for classifiers. The framework is designed to improve detection performance while delivering stable convergence results and decreasing processing time. The proposed design works well in real-world intrusion detection environments.

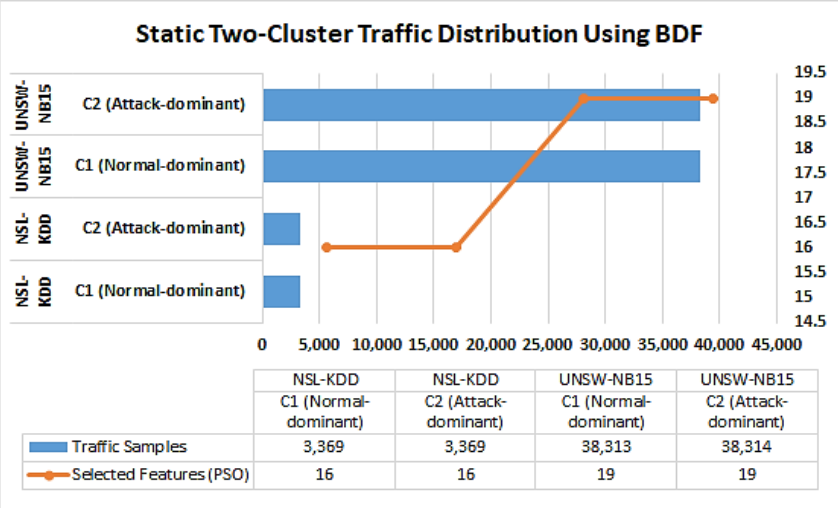


Figure 7. Static two-cluster traffic distribution and PSO-selected feature counts obtained after DBF-based clustering on two IDS datasets.

IV.viii. Cross-Validation and Robustness Analysis for Proposed Approach

The DBF–PSO system, which claims high accuracy, needs more testing to determine whether its outcomes are trustworthy and not the result of excessive model training. For this purpose, 10-fold cross-validation was conducted on both the NSL-KDD and UNSW-NB15 datasets. The study used a method that split each dataset into ten parts, which resulted in nine parts being used for training and one part being used to validate the results. The process was repeated ten times so that each subset was used once for validation, and the final performance was calculated as the average across all folds. The assessment was conducted using two types of network traffic conditions, which included both balanced and original dataset traffic patterns, which showed different traffic distribution levels. In the balanced scenario, the number of samples in normal and attack classes was adjusted to achieve equal representation, whereas the imbalanced scenario maintained the original dataset distribution to reflect realistic network traffic conditions.

The cross-validation results summarized in Table 10 show stable performance across both datasets. The proposed framework achieved 98.92% accuracy with an F1-score of 98.87% under balanced traffic conditions for the NSL-KDD dataset, whereas the model achieved 98.74% accuracy with 98.69% F1-score under the imbalanced scenario. The UNSW-NB15 dataset showed that the framework achieved 99.34% accuracy with

99.29% F1-score under balanced conditions and 99.17% accuracy with 99.12% F1-score under imbalanced conditions. Used a paired t-test statistical analysis to determine how stable their results were. The analysis produced p-values lower than 0.05, which confirmed that the DBF–PSO framework achieved performance improvements that reached statistical significance. The results demonstrate that the proposed model maintains consistent performance across different validation folds and traffic distributions, which shows strong robustness while decreasing overfitting risk.

Table 10. Validation results of the proposed DBF–PSO framework using 10-fold cross-validation under balanced and imbalanced traffic scenarios.

Dataset	Traffic Condition	Validation Method	Accuracy (%)	F1-Score (%)	p-value
NSL-KDD	Balanced	10-Fold Cross Validation	98.92	98.87	<0.05
NSL-KDD	Imbalanced (Original)	10-Fold Cross Validation	98.74	98.69	<0.05
UNSW-NB15	Balanced	10-Fold Cross Validation	99.34	99.29	<0.05
UNSW-NB15	Imbalanced (Original)	10-Fold Cross Validation	99.17	99.12	<0.05

IV.ix. Assessment of the Proposed Model Against Existing Techniques

The proposed PSO + DBF intrusion detection model has undergone performance testing against multiple existing IDS models, which use different machine-learning or deep-learning classifiers along with feature selection algorithms, as shown in Table 11. Earlier works have utilized meta-heuristic optimization techniques like PSO, GWO, GA, SSA, DE, and CS to reduce the dimensionality of feature space with classifiers such as SVM and RF, Deep Neural Networks, LSTM networks, and XGBoost. The hybrid techniques achieved good detection performance but required either extensive feature sets or deep-learning architectures, which needed high computational power, while their performance varied between different datasets. The proposed PSO + DBF model achieves high detection accuracy through its need for fewer features, which makes it better suited for real-time intrusion detection than other operational modes. The NSL-KDD dataset showed that the proposed method could achieve 99.36% detection accuracy with 16 optimized features, which outperformed PSO + PCA + SVM (98.50%) and K-Means + RF + CS (98.70%), and LSTM + SSA (97.89%). The proposed model achieved 99.89% accuracy on the UNSW-NB15 dataset with 19 features, which exceeded the performance of PSO + XGBoost (99.51%) and ResNet–BiGRU + GA + PSO (95.17%), and DNN + PSO (82.24%) while approaching the results of DT + RF (99.86%).

Table 11. Comparison of proposed DBF-based clustering on PSO-optimized classification with recent IDS approaches.

Study	Name of IDS Dataset	Proposed Approach	Count of features used	Accuracy	Recall	F1-score	Precision
[XXVIII]	NSL-KDD	DT + RF	Not reported	99.86	99.79	N/A	N/A
	UNSW-NB15	DT + RF	Not reported	96.01	97.89	N/A	N/A
[XIV]	UNSW-NB15	PSO + XGBoost	Not reported	99.51	99.36	99.32	99.3
[XXX]	UNSW-NB15	ResNet-BiGRU + GA + PSO	Not reported	95.17	93.92	94.27	94.63
[XXVII]	NSL-KDD	PSO + PCA + SVM	18	98.5	96.9	98.4	86.9
[XXXI]	UNSW-NB15	Deep Neural Network + PSO	Not reported	82.44	81.06	85.94	77.68
	NSL-KDD	Deep Neural Network + PSO	Not reported	90.42	88.75	85.24	93.66
[XXIX]	NSL-KDD	CS + RF + K-means	19	98.7	99.07	98.68	98.28
	UNSW-NB15	CS + RF + K-means	22	99.78	99.42	99.59	99.77
[XI]	NSL-KDD	LSTM + SSA	Not reported	97.89	100	97.73	91.67
[XII]	UNSW-NB15	SS + RF	Not reported	95.45	95.2	95.83	96.66
	NSL-KDD	FPA + RF	Not reported	98.79	99	98.94	98.95
Proposed	**NSL-KDD**	**DBF + PSO + KNN**	**16**	**99.36**	**99.38**	**99.36**	**99.35**
	UNSW-NB15	**DBF + PSO + KNN**	**19**	**99.89**	**99.84**	**99.84**	**99.85**

V. Conclusion

A PSO-DBF intrusion detection paradigm is introduced in this study, incorporating PSO for adaptive feature selection with DBF-based clustering for structuring network traffic prior to classification. The model was developed to remedy certain key deficiencies noted in the recent IDS solutions that typically encountered problems with feature redundancy, variable detection abilities across datasets, and high computation overhead running into real-time monitoring scenarios. The existing traditional models, which include SVMs, NB, and QDA, and deep neural architectures, use decision boundaries that remain unchanged, need heavy training resources, and show detection performance problems when traffic patterns change. The PSO-DBF method shows significant improvement in classification accuracy on NSL-KDD and UNSW-NB15 datasets with 99.36% and 99.89% precision through the reduction of the feature set to 16 and 19 attributes. The results show that PSO-DBF outperforms many hybrid IDS systems, which use GWO, GA, SSA, DE, XGBoost, RF, DNN, and LSTM

techniques as their base. The timing evaluation results show that the optimized model operates with low detection latency because its features were improved for compatibility with real-time intrusion detection systems. The statistical tests confirmed that the performance enhancements achieved in this study could be replicated by other researchers. The PSO-DBF framework achieves detection accuracy while keeping computational costs low, which enables it to generalize across different intelligent network challenges faced at present. The future research will focus on distributed and incremental learning systems that can handle high-volume streaming data that comes from IoT and edge-cloud security infrastructure. The system will undergo research to understand its ability to withstand adversarial attacks while developing more robust and protective measures. The evaluation of the proposed framework will use recent intrusion detection datasets, which include CIC-IDS2017 and CSE-CIC-IDS2018, along with CPU, memory, and energy efficiency assessment of its computational resource requirements. The PSO-DBF framework serves as a strong candidate for implementation in real-world network-defined systems, which include enterprise security gateways and cloud data-center monitoring platforms. The system provides high accuracy and reduced feature complexity together with efficient inference for real-time IDS/IPS applications, which require precise and rapid performance.

Conflicts of interest

The authors declare no conflict of interest.

References

I. Al-Alyawy, M., Hinckley, S., Mezher, M. H., Husain, S. O., & Al-Fatlawi, A. H. (2024, November). Thermodynamics-based passive house. In AIP Conference Proceedings (Vol. 3229, No. 1, p. 070003). AIP Publishing LLC.

II. Albu-Salih, A. T., Jumaah, M. Y., Al-Fatlawi, A. H., & Najm, H. (2025). Efficient Hybrid Feature Engineering and Supervised Learning Approach for Network Traffic Classification in Intrusion Detection Systems. International Journal of Intelligent Engineering & Systems, 18(6).

III. Aighuraibawi, A. H. B., Manickam, S., Alyasseri, Z. A. A., Abdullah, R., Khallel, A., Al Ogaili, R. R. N., ... & Yahya, A. E. (2024). Hybridizing flower pollination algorithm with particle swarm optimization for enhancing the performance of IPv6 intrusion detection system. Alexandria Engineering Journal, 104, 504-514.

IV. Alzamili, S. L., Baawi, S. S., Kadhim, M. N., Al-Shammary, D., & Ibaida, A. (2025). Efficient feature selection based on Ruzicka similarity for EEG diagnosis. International Journal of Information Technology, 1-15.

V. Abdulkhudhur, S. M., Abboud, S. M., Najim, A. H., Kadhim, M. N., & Ahmed, A. A. (2025). A Hybrid Deep Belief Cascade-Neuro Fuzzy Approach for Real-Time Health Anomaly Detection in 5G-Enabled IoT Medical Networks. International Journal of Intelligent Engineering & Systems, 18(5).

VI. Alrammahi, A. A. H., Sari, F. A. O., Muhammad, Z. A., Kadhim, M. N., Al-Shammary, D., & Ibaida, A. (2025). Enhancing spam detection with advanced feature extraction and unsupervised clustering. International Journal of Information Technology, 1-11.

VII. Alzamili, S. L., Baawi, S. S., Kadhim, M. N., Al-Shammary, D., Ibaida, A., & Ahmed, K. (2026). Ruzicka Similarity-based Brain EEG Clustering for Improved Intelligent Epilepsy Diagnosis. Computer Methods and Programs in Biomedicine Update, 100229.

VIII. Bosso, L., Smeraldo, S., Rapuzzi, P., Sama, G., Garonna, A. P., & Russo, D. (2018). Nature protection areas of Europe are insufficient to preserve the threatened beetle Rosalia alpina (Coleoptera: Cerambycidae): evidence from species distribution models and conservation gap analysis. Ecological Entomology, 43(2), 192-203.

IX. Baawi, S. S., Kadhim, M. N., & Al-Shammary, D. (2025). Efficient clustering approach based on Gower distance for high-dimensional medical datasets. Cluster Computing, 28(12), 756.

X. Baawi, S. S., Kadhim, M. N., & Al-Shammary, D. (2025). Efficient feature selection based on Gower distance for breast cancer diagnosis. Journal of Electronic Science and Technology, 23(2), 100315.

XI. Dash, N., Chakravarty, S., Rath, A. K., Giri, N. C., AboRas, K. M., & Gowtham, N. (2025). An optimized LSTM-based deep learning model for anomaly network intrusion detection. Scientific Reports, 15(1), 1554.

XII. Emirmahmutoğlu, E., & Atay, Y. (2025). A feature selection-driven machine learning framework for anomaly-based intrusion detection systems. Peer-to-Peer Networking and Applications, 18(3), 1-28.

XIII. Hammood, D. A., Alzayadi, L. H. M., Mahmoud, M. S., & Abd Zaid, M. M. (2025). Efficient Hybrid Intrusion Detection Approach based on BPR-GWO for Network Traffic Classification and Improved Network Security. International Journal of Intelligent Engineering & Systems, 18(8).

XIV. Hammood, D. A. (2024, October). A hybrid system based on machine learning and PSO for network intrusion detection. In AIP Conference Proceedings (Vol. 3232, No. 1, p. 020041). AIP Publishing LLC.

XV. Hamad, A. R., Baraa Alsabti, S. M., Najim, A. H., & Kadhim, M. N. (2025). A Hybrid Feature Selection and Machine Learning Approach for Parkinson's Disease Detection from Voice Signals in IoT-Enabled 6G Networks. International Journal of Intelligent Engineering & Systems, 18(5).

XVI. Hashim Albohayah, Z. H., Abed, S. B., Mahdi, A. J., Kadhim, M. N., & Najim, A. H. (2025). Ch-PSO: A Novel Embedded Method based on PSO and Chebyshev Distance for Enhanced Epileptic Seizure Classification Using EEG Brain Signals. International Journal of Intelligent Engineering & Systems, 18(5).

XVII. Jabier, E., Marhoon, A. F., Aldair, A. A., Kadhim, M. N., Al-Shammary, D., & Ibaida, A. (2025). Efficient Kulczynski EEG feature selection for autism spectrum disorder diagnosis over fog and cloud computing. International Journal of Information Technology, 1-17.

XVIII. Kurdi, W. H. M., Rassool, H. A., & Al-fatlawi, A. H. (2021). Evaluation patterns and algorithm for cancer identifications using dynamic clustering. Periodicals of Engineering and Natural Sciences (PEN), 9(2), 462-470.

XIX. Kadhim, M. N., Mutlag, A. H., Hammood, D. A., & Ismail, N. B. H. (2025). Identification of Vehicle Logos in Deep Learning: A Comprehensive Survey. Journal of Techniques, 7(1), 37-47.

XX. Latif, S., Boulila, W., Koubaa, A., Zou, Z., & Ahmad, J. (2024). Dtl-ids: An optimized intrusion detection framework using deep transfer learning and genetic algorithm. Journal of Network and Computer Applications, 221, 103784.

XXI. Malik, R. Q., Alsharfa, R. M., Mohammed, B. K., Al-Fatlawi, A. H., Abd Al-Ameer, M. S., & Najm, H. (2025). A Novel Taneja Distance-based Classifier with PSO-Optimized Feature Selection for Efficient 5G Network Slicing. International Journal of Intelligent Engineering & Systems, 18(6).

XXII. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set", presented at the 2009 IEEE symposium on computational intelligence for security and defense applications, Ieee, pp. 1–6, 2009.

XXIII. Mohammed, M. H., Kadhim, M. N., Al-Shammary, D., & Ibaida, A. (2025). Novel Voice Signal Segmentation Based on Clark Distance to Improve Intelligent Parkinson Disease Detection. Journal of Voice.

XXIV. Mohammed, M. H., Kadhim, M. N., Al-Shammary, D., & Ibaida, A. (2025). EEG-Based Emotion Detection Using Roberts Similarity and PSO Feature Selection. IEEE Access.

XXV. N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set) ", presented at the 2015 military communications and information systems conference (MilCIS), IEEE, pp. 1–6, 2015.

XXVI. Rfys, R. R., Al-Shammary, D., Kadhim, M. N., & Ibaida, A. (2026). Novel ECG Signal Classification based on Minkowski Distance to Enhance Intelligent Arrhythmia Detection Systems. Smart Health, 100645.

XXVII. Raghunath, M. P., Deshmukh, S., Chaudhari, P., Bangare, S. L., Kasat, K., Awasthy, M., ... & Waghulde, R. R. (2025). PCA and PSO based optimized support vector machine for efficient intrusion detection in internet of things. Measurement: Sensors, 37, 101806.

XXVIII. Umar, M. A., Chen, Z., Shuaib, K., & Liu, Y. (2025). Effects of feature selection and normalization on network intrusion detection. Data Science and Management, 8(1), 23-39.

XXIX. W. H. Madhloom Kurdi, I. A. Alzuabidi, A. H. Najim, M. N. Kadhim, and A. A. Ahmed, "Efficient Two-Stage Intrusion Detection System Based on Hybrid Feature Selection Techniques and Machine Learning Classifiers", International Journal of Intelligent Engineering & Systems, Vol. 18, No. 3, 2025.

XXX. Xia, Z., He, S., Liu, C., Liu, Y., Yang, X., & Bu, H. (2024). PSO-GA Hyperparameter Optimized ResNet-BiGRU Based Intrusion Detection Method. IEEE Access.

XXXI. Yılmaz, A. A. (2025). A novel deep learning-based framework with particle swarm optimisation for intrusion detection in computer networks. PloS one, 20(2), e0316253.

XXXII. Y. S. Mezaal, "New compact microstrip patch antennas: Design and simulation results," *Indian J. Sci. Technol.*, vol. 9, no. 12, 2016. 10.17485/ijst/2016/v9i12/85950