Feature Selection Techniques in Intrusion Detection Systems: A Review

Ahmad Salim[1], Obaid Salim[2], Omar Muthanna Khudhur[3,*], Shokhan M. Al-Barzinji[4], Farah Maath Jasem[5]

[1]Middle Technical University, Iraq

[2]General Directorate of Education Anbar, 31001, Iraq

[3]Department of Computer Engineering Techniques, College of Technical Engineering, University of Al Maarif, Al Anbar, 31001, Iraq

[4]Department of Computer Networks Systems, College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq

[5]College of Computer Science and Information Technology, University of Anbar, Anbar, Iraq

Emails: ahmadsalim@mtu.edu.iq; multiknowlge@gmail.com; omar.m.khudhur@uoa.edu.iq; shokhan.albarzinji@uoanbar.edu.iq; Farahmaath86@uoanbar.edu.iq

Abstract

Intrusion detection has garnered significant attention as researchers strive to develop sophisticated models characterized by their high accuracy levels. However, the persistent challenge lies in creating reliable and effective intrusion detection systems capable of managing vast datasets under dynamic, real-time conditions. The effectiveness of such systems largely depends on the chosen detection methodologies, specifically the feature selection processes and the application of machine learning techniques. This paper offers a comprehensive review of feature selection methods employed in the realm of intrusion detection research. It examines various dimensionality reduction strategies, followed by a systematic classification of feature selection techniques to assess their impact on the training phase and subsequent detection efficacy. The focus was on the wrapper, filter feature selection methods, where the methods used were analysed, and their strengths and weaknesses were revealed. The identification and selection of the most pertinent features have been shown to significantly enhance the detection performance, not only in terms of accuracy but also in reducing computational demands, underscoring its critical importance in the architecture of intrusion detection systems.

Keywords: Network security; Intrusion detection; Machine learning; Feature selection; Wrapper; Filter

1. Introduction

The most contemporary and the dominant feature of the global information ecosystem is nothing but the explosive growth of data driven by the rise of Internet and technology. This results in the ever-increasing amount of data that is being produced and this boost in data generation can be attributed to many factors like everyone having a mobile now, IoT gaining popularity or cloud computing becoming mainstream. As a result, consequences for network and information security are vast and varied [1]. Nevertheless, the presence of anti-attack systems such as firewalls and intrusion detection systems (IDS) is needed to prevent and protect against network attacks, keeping secure borders 2 with a continuous monitor to detect threats and act effectively in response.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

IDSs are central to the network security architectures maintaining data integrity, confidentiality, and availability. These systems are also very useful in monitoring network traffic, which keeps ongoing 24/7 and watches network environment for unauthorized access attempts and attacks [4]. There are two types of IDS; one of them is HIDS (Host-based Intrusion Detection Systems), and the other one is NIDS (Network-based Intrusion Detection Systems) HIDS are installed on every host or device and will be monitoring the incoming and outgoing packets from that specific host, giving a deep-dive into activities at a host level. Whereas NIDSs are strategically placed throughout the network, this allows them to listen to all incoming and outgoing traffic expected from all devices on the network, providing a wider glimpse at how healthy or safe (or not) the network is [5]. Besides, IDS are divided into signature-based as well as anomaly-based on a wider scale. Like with antivirus systems, a signature-based IDS uses predefined patterns of known threats that it recognizes and responds to declared malicious behaviours. Anomaly-based IDS however, use machine learning and statistical models to create a baseline for normal traffic and detect a deviation from this behaviour which could indicate security breaches [6][7].

Intrusion detection systems (IDSs) are one of the main areas for applying machine learning (ML) techniques, as these use algorithms to analyse network data looking for patterns that could be seen as security break-ins. In ML models, feature selection (FS) is crucial for improving performance, interpretability, preventing overfitting, avoiding the curse of dimensionality and speeding up training by focusing on the most relevant features in prediction tasks [8][9]. FS techniques in ML is broadly categorized into two types, based on the presence and use of labels in data. Supervised methods are based on the labels and used to select significant features, which can be divided into three types: wrapper methods, filter methods and embedded methods [10][11]. Wrapper methods will try different combinations of features relying on some learning algorithm that uses them as inputs and returns a score it should be able to reach the optimal subsets. Filter methods, in contrast, use statistical measurements to calculate what kind of relationship (kind of correlation or dependence) between input variables and target machine learning algorithm. Embedded methods are embedded within particular algorithms, such as decision trees, which intrinsically perform feature selection as part of their construction process [10][12]. Conversely, unsupervised methods operate without labelled data, focusing instead on analysing the distribution and relationships among features to discern structure and reduce dimensionality without guidance from a target outcome [13][14].

In the scientific literature, several works have been devoted to surveying machine learning algorithms with applications to Internet traffic classification. However, few attempts have been made in the field of IDS [15][16], and two major drawbacks have emerged: (1) FS techniques used in IDS have not been analysed; (2) The field lacks updated analysis the FS techniques used in IDS. In general, not many new algorithms have been compared. Our manuscript aims to address these gaps, and makes the following contributions:

- We reviewed the latest techniques used to select features in IDSs.
- We analysed the important techniques adopted for feature selection and focused on two types of feature selection, namely wrapper and filter.

The rest of the manuscript is organized as follows: Section 2 explains the feature selection techniques and their types. In Section 3, the wrapper and filter techniques used in IDSs are reviewed. Section 4 analyses these techniques. Meanwhile, conclusions and future works in Section 5.

2. Feature Selection

Feature selection is a crucial step in the machine-learning pipeline. It involves selecting a subset of relevant features (variables, predictors) for use in model construction [17]. Feature selection methods help in reducing dimensionality, removing irrelevant or redundant data, improving learning accuracy, and improving model interpretability [11][18]. Machine learning techniques rely heavily on feature selection methods to reduce complexity and enhance model performance. Among the most significant domains that benefit from these methods are cybersecurity models [19]. There are two types of feature selection: supervised and unsupervised. The types of supervised feature selection techniques can be broadly classified into three main categories: filter methods, wrapper methods, embedded methods, and hybrid methods [16][20]. The classification of feature selection techniques shown in fig. 1.

a. Filter Methods

Filter methods evaluate the relevance of features by their intrinsic properties, independent of any machine-learning algorithm. These methods are generally fast and scalable, making them suitable for high-dimensional datasets. Key techniques include [12][21]:

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

- Correlation Coefficient: Features are selected based on their correlation with the target variable. Features that show higher correlation are considered more relevant.
- Chi-Square Test: This test is used to determine the dependence between categorical variables and the target. Features with higher chi-square values relative to the target are selected.
- Information Gain: Measures the reduction in entropy or surprise from transforming a dataset in some way. It is commonly used in training decision trees.
- Variance Threshold: This technique removes features whose variance does not meet a certain threshold. It is based on the premise that features that do not vary much within themselves carry little information.
- ANOVA F-test: Used to select continuous input features for a categorical target.

b. Wrapper Methods

Wrapper methods consider the selection of a set of features as a search problem, where different combinations are prepared, evaluated, and compared to other combinations. A predictive model is used to evaluate a combination of features and assign a score based on model accuracy. Key techniques include [16][22]:

- Recursive Feature Elimination (RFE): Iteratively constructs models and removes the weakest feature (or features) until the specified number of features is reached. This is usually done with a particular type of machine learning algorithm.
- Sequential Feature Selector: This can be a forward selection method where features are sequentially added to an empty set until an optimal feature subset is found or a backward elimination method where features are sequentially removed from a full set.
- Genetic Algorithms (or meta-heuristic algorithms): The process of natural selection inspires these; this approach uses techniques such as mutation, crossover, and selection to generate feature subsets, evaluating them using a fitness function.

c. Embedded Methods

Embedded methods integrate feature selection as a part of the model training process and are specific to given learning algorithms. These methods can be more efficient than wrapper methods since they incorporate feature selection and model training. Key techniques include [12]:

- LASSO (Least Absolute Shrinkage and Selection Operator): LASSO is a regularization technique that includes a penalty term to the loss function proportional to the absolute value of the coefficients. Features with coefficients that shrink to zero are removed from the model.
- Ridge Regression: Although it does not perform feature selection in the traditional sense, (all features remain but some are minimized), it is useful for multicollinearity data.
- Decision Trees and Random Forests: These models inherently perform feature selection by choosing which features to split on at each node during the tree building process. Feature importances can be derived from these models.

d. Hybrid Methods

Hybrid methods combine the qualities of filter and wrapper/embedded methods to form a robust feature selection methodology. These might use a filter method to reduce the dimensionality significantly before a wrapper method is applied to find an optimal subset [20].

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

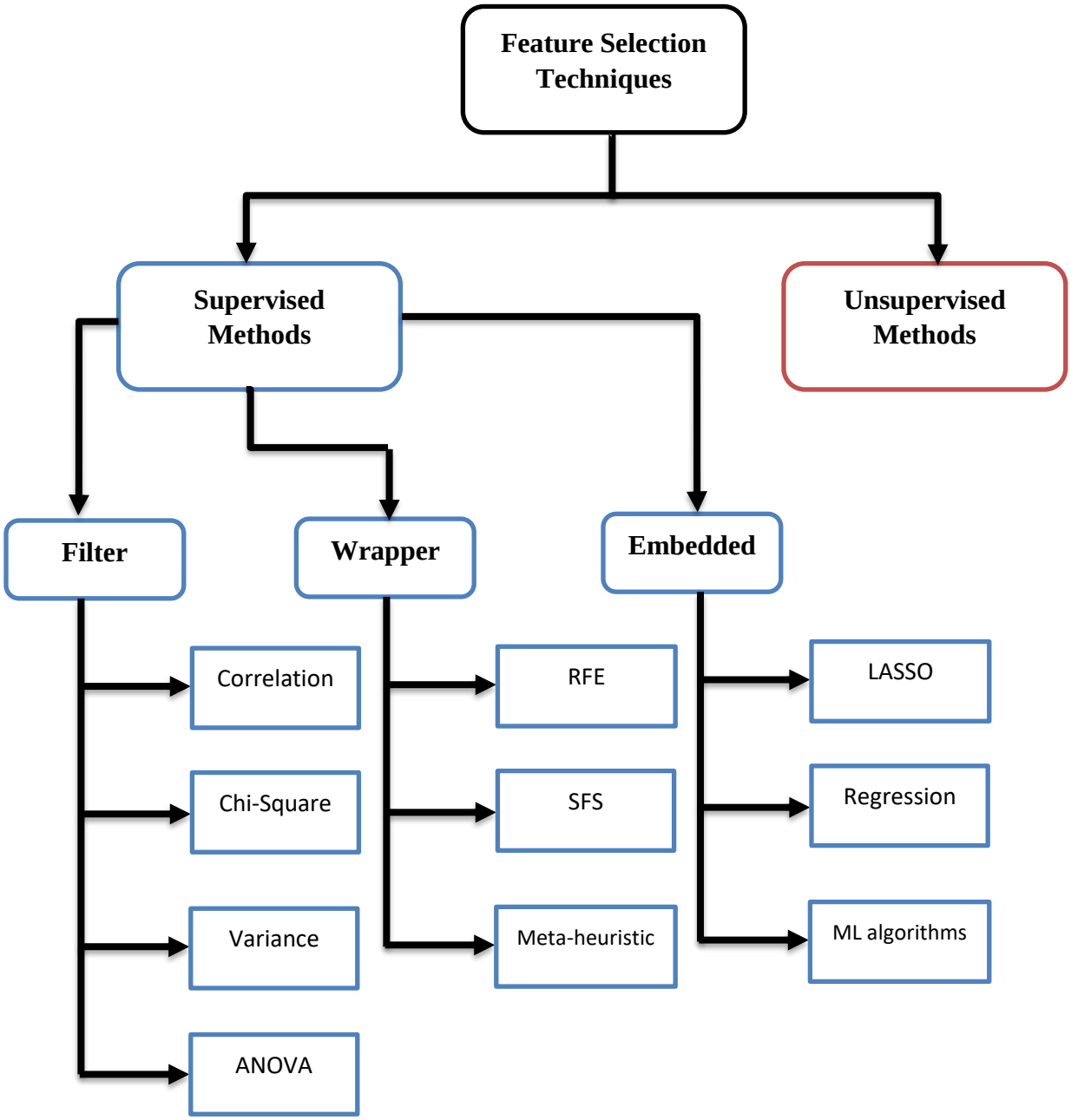

Figure 1. Classification of feature selection techniques

3. Feature Selection Techniques in IDS

This paper explores the critical role of feature selection methods in the field of IDS. By reviewing different techniques and their impact on system efficiency, it aims to illustrate how the selection of refined features enhances both the accuracy and computational efficiency of IDS. The work is based on reviewing the methods for selecting the feature: filter and wrapper.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

3.1. Review of Filter FS Methods in IDS

Filter feature selection methods are fundamental preprocessing steps in data mining and machine learning, providing a powerful means for dimensionality reduction before the application of an algorithm. These techniques determine the importance of features by measuring their inherent properties using statistical metrics like correlation coefficients, chi-squared test, information gain, etc. The strong benefit of filter methods is their quick evaluation of feature subsets, as they are not trained with the model. These are very useful to handle large data sets, which help in reducing the computing time and hence allow us to develop focused models quicker.

In [23], posed a multi metric-based feature selection method for increasing the performance IDS. These different metrics are introduced into the seed of feature selection to predict which features are essential in detecting intrusions and further enhancing the performance of IDS. The paper highlighted the role of feature selection in determining how well an IDS works, and provided a good knowledge about the problems faced by this area. It is important to mention that the flexibility of the algorithm allows us to adapt it easily for different intrusion detection scenarios, increasing its robustness and generalization. Nevertheless, the method is not properly specific with respects to both metrics and summation / weighting mechanisms employed this fact can hamper transparency and reproducibility. Although the algorithm was not without its limitations, it achieved good results on benchmark datasets like KDD'99, NSL-KDD and CDMC2012 with accuracies of 93.88%, 84.04% and 99.20% respectively.

There are numerous researchers used correlation as a feature selection technique, in [24] a Correlation-based feature selection framework was proposed to improve anomaly-based IDS performance. This work proposes a CFS-AE model, which incorporates a correlation-based feature selection method that selects features based on their statistical correlations with an Auto encoder designed to compress and decompress the input data learning effectively the representation of information relevant for anomaly detection. In this way, it works perfectly to filter noise from the data and reduce dimensionality, increasing the ability of the model to be able to identify patterns in such kind of dataset. The experimental results on two benchmark datasets, NSL-KDD and CIC-IDS2017 dataset are reported to demonstrate the effectiveness of the model. The model attained a higher accuracy of 94.32% on the NSL-KDD dataset, 97.71% on the CIC-IDS2017 dataset when compared to conventional method MGT CFS-AE Model offers an effective, albeit blunt instrument to measure acuity but this has limitations. Firstly, the correlation-based approach deals only with linear dependencies, which might not be appropriate if features are less likely to interact in a simple way from functional perspective. During 2020, equally problematic is that this may grossly oversimplify the security landscape. In addition, the model has a very specific threshold condition for feature selection and is susceptible towards outliers that are ubiquitous inside IDS environments.

In [25], introduced a new cross-correlation based feature selection technique (CCFS) and examine its ability to improve the performance of IDSs. CCFS uses statistical measures to evaluate both: i) pairwise relationship (cross-correlation) between features and ii) how relevant each feature is with respect to the target variable. In comparison with two other feature selection approaches, the evolutionary-based cuttlefish algorithm (CFA) and entropy-based mutual information feature selection (MIFS). We evaluate the strengths of these approaches by applying four classifiers: Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT) and K-Nearest Neighbour (KNN). While the CCFS approach appears to be viable, there is an important shortcoming in the preprocessing step of it — simply treating discrete attributes as being copied values. That potentially more efficient encoding approach (like one-hot-encoding) will give a better representation of the categorical data, and therefore have an impact on how good is the feature selection and classification process.

In [26], introduced a new intrusion detection method built on the use of Correlation-based Feature Subset Selection (CFS) to feed an Artificial Neural Network (ANN), thus reducing the number of features used to identify threats and increasing accuracy. The process determines characteristics of network traffic that have high correlation for classifying network utilization either normal or anomalous, and it keep these attributes. This detection can be done in two distinct stages: automatically selecting important features; and processing these selected features through the ANN to detect changes that could indicate intrusions. Correlation-based feature selection is used to increase the ANN learning speed by regularization of input, elimination of noise and irrelevant data points. Our results are encouraging and we find that detection accuracy is superior to traditional IDS implementations with no feature selection. The results are further robust by validating on multiple datasets such as KDD Cup 99 and NSL-KDD. A resulting limitation, discussed in the paper is that only linear correlations are considered (and not nonlinear ones needed to uncover more advance cyber threats). Also, the paper neither fully explore what potential biases may be introduced by the correlation threshold setting nor consider how different network environments might affect this system's performance.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

In [27] introduced a hybrid IDS, unique in that, it overcomes the two challenges experienced by classic of stacking-based methods; (1st) offering a remedy for classification performance especially with high-dimensional features and (2nd), its robustness to the underlying classifier. In order to reduce the complexity of features, We implement CFS-DE algorithm to select the subset features and then design a weighted stacking combining with training results from different classifiers so that enhances its importance when they have good training accuracy, but decreases their influence if they perform worse. The hybrid method improves the efficiency and robustness of classification. Experimental results on NSL-KDD as well as CSE-CIC-IDS2018 illustrate that the proposed method could achieve 87.44% accuracy 89.09% precision, 86.61 %recall and FS score of 88.25% & KDDTest+, accuracy. Finally, on it achieves ultra-highs accuracy/precision/recall/view all topics <99%. In CSE-CICISC2018, the proposed method achieves 99.25%, 100% and 99.88% TPR, TNR and F1 score respectively. Comparison analysis results that the proposed CFS-DE weighted stacking type of IDS are better than traditional machine learning models and existing framework as well. This also concludes that it indeed has a powerful classification performance in intrusion detection. The problems detected include: No technique has been used to show the importance of features, Value of correlation measure among features is not appearing anywhere and there is incomplete description for understanding the pre-processing steps.

Numerous studies have used other statistical techniques as for feature selection in IDS. ANOVA is one of the commonly used statistical measures for feature selection; Venkatesan [28] designed IDS utilizing feature selection techniques based on machine learning (ML) algorithms, specifically ANOVA F-Test, with the NSL-KDD dataset. By employing SVM, random forests, and decision trees, the study seeks to identify pertinent features for intrusion detection. However, the obtained results, displaying an accuracy of 0.99, are deemed unrealistic, prompting a critical examination of the ANOVA F-Test method. The identified drawbacks include the assumption of equal variance, sensitivity to outliers, challenges in handling nonlinear relationships, and the neglect of feature interactions. These limitations underscore the need for alternative feature selection methods to enhance the IDS's accuracy and robustness. In [29], the authors delve into the world of IDS by evaluate feature selection techniques. Specifically, they explored the effectiveness of two techniques: ANOVA F-test and sequential feature selection (SFS). The datasets under scrutiny include NSLKDD, Kyoto2006, and UNSW_NB15. The proposed approach consists of two distinct steps. First, the ANOVA F test assigns scores to features based on their relevance to the label. These scores serve as the basis for feature classification. Subsequently, the SFS method is used to determine the optimal feature set. It is worth noting that the performance gains from this approach are significant: multi-class problems see an average improvement of more than 10%, while binary classification benefits by approximately 5%. Experimental results across different datasets are as follows: NSL-KDD achieved an impressive 99.45% accuracy, Kyoto2006 showed strong performance with 97.42% accuracy, and UNSW-NB15 achieved 70.63% accuracy. However, there are some limitations. First, the method is sensitive to multicollinearity between features. Second, there is a risk of over-equipping. Finally, incremental exploration of the feature space in SFS may lead to overlooking globally optimal feature sets, especially in complex search spaces. In addition, a set of statistical methods were combined and used together to improve the feature selection in IDS. In [30], suggested a fusion approach to enhancing feature selection in IDS. The study integrated statistical importance measures to improve the interpretability and effectiveness of feature selection, significantly boosting detection accuracy and reducing false positives. However, we noted several weaknesses, including high computational costs associated with calculating statistical importance for large datasets, potential overfitting in scenarios with imbalanced data, and the necessity for frequent updates to adapt to evolving threats. Additionally, the method's real-time applicability was questioned due to its resource-intensive nature. Despite these challenges, the paper made substantial contributions to IDS research, demonstrating the potential of combining statistical importance with deep learning to enhance IDS performance.

In [31], a feature selection method for intrusion detection system in Internet of Things (IoT) environment using information gain (IG) and gain ratio (GR) is presented to choose top 50% ranked features. The proposed method is tested on the BoTIoT dataset and benchmarked against KDD Cup 1999. Results It illustrates better performance by maximum score with classifier filter particularly for JRip classifier with accuracy and less time for constructing the model. In particular, the performance of our system is compared with the existing methods on two datasets, further demonstrating its superiority. Of course, there are disadvantages that come with the information gain and gain ratio. Everything from biases toward more classes to being sensitive to noise, ignoring feature interactions and the computational complexity of some methods were addressed, as are some potential anomalies in intrinsic information measures. Although the efficacy of the proposed feature selection is far from perfect, it can be considered an efficient one for improving IDS performance in IoT networks.

Notably, [32] viewed at this favourably and you uploaded in the last… proposed a new way of improving feature selection effectiveness and accuracy of IDS. This research undoubtedly enhanced the performance of IDS using a well-

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

efficient feature selection and dimensionality reduction approach. However, the reviews also reflected some limitations of SR: its scalability (on large data), significant computational time-cost overhead that could prevent it from being applied in real-time, and difficulties to deploy SR practically under dynamic network conditions. These concerns notwithstanding, the paper was singled out for raising the bar in IDS methods (albeit over a decade ago), with further research stipulated to tackle this limitation and make the method usable in more real-world scenarios.

Furthermore, [33] presents a novel solution for feature selection for intrusion detection systems: applying multiple filter-based algorithms together in order to increase the performance. This method focuses on reducing dimensions effectively and keeping detection accuracy high, especially for network traffic analysis since we work with a large number of dimensions. Yet, there were several limitations- it introduced extra computational overhead by maintaining and integrating a wide number of filters in the overall system, that can be operationally challenging to scale up even for medium-scale networks (say, tens thousands of devices), thus struggling with large network deployments scenarios. It also comes with slower real-time processing times because you might need few milliseconds to detect an event and then couple more milliseconds to construct a response causing delay at bestinance actions). Furthermore, the challenge of ensuring up-to-date components with a suite required regular maintenance and updating to adapt to ever-changing threats was pronounced as problematic. In Table 1, we present an overview of the most relevant related works in the area of filter feature selection approaches. A summary of the most relevant feature selection methods is shown in Table.

Table 1: Comparison of filter FS techniques in IDS.

Ref.	FS techniques	Datasets	Accuracy	Weakness
[23]	· Information Gain · Gain Ratio · Symmetrical Uncertainty · Chi-Square (χ^2) · ReliefF	- KDD'99 - NSL-KDD - CDMC2012	99.20%	· Lack of specific details about metrics. · Combination and weighting problem. · Absence of empirical evaluations.
[24]	· Correlation-Based Feature Selection and Autoencoder	- NSL-KDD - CIC-IDS2017	97.71%	· Linear relationships only. · Oversimplification. · Threshold dependency. · Impact of outliers.
[25]	· Cross-correlation	- KDD Cup 99 - NSL-KDD - CIC-IDS2017	97.91%	· Numeric values were used for categorical attributes. · One-hot encoding was not considered. · This approach was less suitable for symbolic data.
[26]	· Correlation · Neural network	- NSL-KDD - UNSW-NB	97.49%	· High computational complexity and processing time. · Less effective in identifying novel attacks. · Challenges in managing false alarms in dynamic network environments.
[27]	· CFS-DE	- NSL-KDD	99.87%	· Missing feature correlation values.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

	· Weighted stacking algorithm	- CSE-CIC-IDS2018		· Incomplete description of preprocessing steps.
[28]	· ANOVA F-Test	- NSL-KDD	99%	· Equal variance assumption. · Sensitivity to outliers. · Difficulty in handling non-linear relationships. · Does not consider feature interactions.
[29]	· ANOVA F-test · Sequential Feature Selection	- NSLKDD - Kyoto2006 - UNSW_NB15	99.90%	Sensitivity to multicollinearity. Risk of overfitting. Limited Exploration of Feature Space
[31]	· Information Gain · Gain Ratio	- KDD Cup 1999	99.99%	Bias toward features with more categories Sensitivity to noise Intrinsic information anomalies

3.2. Review of Wrapper FS Methods in IDS

Wrapper feature selection techniques evaluate subsets of features by repeatedly training and testing the model. These methods evaluate subsets of features based on their predictive performance using a specific learning algorithm. Unlike filtering methods, wrapping techniques take into account feature interactions and model complexity during selection. Examples include recursive feature elimination (RFE) and sequential feature selection algorithms. Although computationally intensive, encapsulation methods often produce more accurate subsets of features tailored to the chosen learning algorithm [34]. Wrapper feature selection techniques use optimization algorithms to search for an optimal subset of features that maximize model performance. These algorithms repeatedly evaluate different feature sets by training and testing a predictive model. Commonly used optimization algorithms such as: genetic algorithm (GA), particle swarm optimization, simulated annealing, and hill climbing. By exploring the feature space and considering interactions between features, these approaches aim to identify the most relevant subset to improve model accuracy and generalization [35][36].

Extreme Learning Machine (ELM) has contributed significantly to the selection of the best features in IDS. In [37], received considerable attention for its innovative approach to enhancing IDS. The study combined Differential Evolution (DE) with ELM as a wrapper feature selection method to improve IDS performance. DE was used to identify the most relevant features, while ELM provided fast and accurate evaluation of these features. The method demonstrated significant improvements in detection accuracy and efficiency, surpassing traditional approaches. It was rigorously tested on benchmark datasets, showing enhanced detection rates and reduced false positives. However, the approach introduced computational complexity and potential overfitting risks. Almasoudy [38] presented a new model for selecting features for IDS using differential evolution technique. The feature selection method is mainly based on GA Using 41 features from the NSL-KDD dataset from which only 7 features were selected, the proposed method achieves an accuracy of 80.15% and 87.53% for five- and two-classes, respectively. It addresses the shortcomings of signature and anomaly-based intrusion detection methods by efficiently selecting relevant features and using ELM classification. Additionally, it contributes to enhancing detection rates and reducing false alarms compared to existing methods. However, limitations include reliance on a single dataset and a fixed number of selected features, making the technique less universally applicable across datasets. Despite drawbacks such as computational complexity and stochastic nature, genetic algorithms remain valuable for navigating complex search spaces in feature selection tasks. In other studies, GA was also adopted for feature selection in IDS, [39] introduced a hybrid wrapper feature selection

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

method for intrusion detection, combining a GA and ELM optimized with SVM classifier. Addressing the challenge of parameter selection in extreme learning machine, the GA optimizes its weights to enhance performance. Subsequently, the optimized ELM serves as an estimator in sequential forward selection to identify key features. The proposed method achieves impressive accuracies of 99% and 86% for the IoT_ToN network and UNSWNB15 datasets, respectively. However, employing a GA for ELM weight optimization entails potential limitations including computational complexity, risk of overfitting, and the need for parameter tuning. In [40], they investigated the development of IDS, which employed hybrid classifiers combined with meta-heuristic algorithms. In order to address this, the study was based on the using GA and Grey Wolf Optimization (GWO) as a way of selecting appropriate features, increasing detection percentage by ignoring irrelevant data. Once the hybrid classifiers are used, the IDS works better as a whole, so that more network attacks could be identified and classified. The study also pointed out NOA's improvement when it comes to recognizing various kinds of cyber-attacks.

Correspondingly, swarm optimization algorithms have contributed well to the field of feature selection in machine learning systems. In [41] investigated the efficacy of feature selection using a random forest algorithm to streamline intrusion detection, followed by a comparative analysis of various classifiers, including k Nearest Neighbour, SVM, Logistic Regression, decision tree, and Naive Bayes, to evaluate IDS performance metrics. Employing particle swarm optimization (PSO) on selective features from the NSL-KDD dataset, the research yields promising results with low computational complexity, achieving 99.32% efficiency and a 99.26% detection rate using a subset of 10 features from the original 41. Despite achieving a good accuracy but limitations, include reliance on a single dataset and a fixed number of selected features, underscoring the need for validation across diverse datasets and exploring variable number of selected features.

Alazzam [42] introduced a feature selection method for IDS employing the Pigeon Inspired Optimizer (PIO) to streamline the selection process. Utilizing three widely used datasets, namely KDDCUP 99, NLS-KDD, and UNSW-NB15, the study proposes two distinct methods: Sigmoid_PIO and Cosine_PIO, with the latter demonstrating superior performance. Notably, the Cosine_PIO method achieved high accuracy scores of 0.96 for the KDDCUP 99 dataset. However, the research identifies certain limitations, particularly related to fitness computation and associated weights, as well as computational cost considerations.

The novel to enhance IDS tailored for cloud computing environment has been proposed in [43]. The technique employed an algorithmically driven optimization mechanism of feature selection to assist IDS in the selection to reduce data dimensionality and thus aid in improving the accuracy and efficiency. Experimental results validated the proposed method on well-known datasets, serving as a useful benchmark to compare with existing methods. Strengths of the study: (1) comprehensive feature selection, (2) better computational efficiency, (3) robust validation and (4) reproducibility in diverse cloud environments. However, problems existed such as the complex algorithm implementation & computational consumption and possible scalability issue coupled with high potential for over-fitting and dependence on validation datasets quality.

Has been shown as an improved IDS created for cloud area in [44]. Imbalanced data: SMOTE (synthetic minority over-sampling technique) was applied. This paper has developed an innovative hybrid feature selection technique by involving Information Gain (IG), Chi-Square (CS) and Particle Swarm Optimization (PSO) where IG & CS have been utilized as fitness functions for PSO. The Random Forest (RF) model was applied for attack detection and classification. Supported by high frequent features importances, the system was validated by both UNSW-NB15 and Kyoto datasets with over 98.39% accuracy on multi-class classification scenario for UNSW or Kyoto benight class, and over 99.25% for the other class respectively; Although this is an encouraging finding regarding the impractical promising method that requires more challenging balancing datasets balancing and may introduce over-fitting. Overall, the research significantly enriched improving the cloud IDS over at tonging novel techniques in feature selection and treating data imbalance.

The experimental setup was motivated by the study in [45], placed upon the use of feature selection and data normalization techniques to improve NIDS performance. Experimental results indicated that different feature selection and normalization methods yield significant differences in terms of accuracy, detection rate, and efficiency for NIDS by the comparative analysis used in the study. To determine the most relevant features from the dataset, this work employed different feature selection strategy including Principal Component Analysis (PCA) (refer to [1]) for dimensionality reduction, Information Gain (IG) and Chi-Square (CS). Moreover, standard min-max scaling, z-score normalization and Decimal Scaling were used to normalize the data. In most cases, the NIDS performs well but evaluation using machine learning classifiers (i.e., Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors(kNN) in KDD Cup 99 and NSL-KDD standard datasets. The results showed that the detection accuracy of

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

NIDS was significantly enhanced with feature selection and its computational overhead was notably reduced. PCA and IG yielded better results among feature selection techniques. Normalization boosted the performance of the system, while min-max scaling produced more accuracy and detection rate. On the other side, multiple challenges were identified in the study – among them -the difficulty of combining more than one method and the risk of analysing large datasets in general for overfitting.

In [46] tackled the problem that high-dimensional intrusion-detection datasets slow classifiers and inflate false alarms, proposing a two-stage feature-selection scheme called GIGA. In the first stage, the authors compute each feature's Gini impurity (GI) and discard those with low importance; in the second stage, they run a genetic algorithm (GA) with a decision-tree-based fitness function to identify an optimal subset of features. They evaluate this scheme on three benchmark datasets—CIC-IDS2017, CSE-CIC-IDS2018 and CIC-DDoS2019—and report that it slashes the feature space from 71 to 8, 70 to 4 and 69 to 8 attributes respectively. Despite the dramatic dimensionality reduction, test accuracies increase slightly from 99.31 % to 99.52 %, 96.01 % to 97.19 % and 97.95 % to 99.98 %. The GIGA-selected feature sets improve performance across different classifiers such as Random Forest and Decision Tree by boosting detection accuracy and reducing both false positives and false negatives. Strengths of this work include the substantial reduction in feature dimensionality and the corresponding drop in false alarms without sacrificing accuracy. Potential limitations are that the gains in accuracy are relatively modest given the already high baseline and that genetic algorithms can be computationally expensive. In [47] the wrapper-based selection approach was adopted to reduce the feature space, employing SVM, Random Forest and Naïve Bayes to identify the 20 most informative features, a step that raises accuracy from 90 % to 93 %. The selected features are fed into a transformer encoder augmented with multi-head attention, global average pooling and a dropout-regularized multilayer perceptron; this architecture enables the model to learn complex feature relationships and achieve a detection score of 0.92 versus 0.87 without multi-head attention. On the UNSW-NB15 dataset the complete system attains 93%, outperforming contemporary CNN and other transformer-based models. While its integration of feature selection, (SMOTE) and attention mechanisms represented a clear strength and illustrates how pre-processing and model design can boost detection performance, the approach is computationally heavy, tested on a single dataset and less interpretable due to the transformer's complexity. Table 2 shows a summary of the most important previous works in the field of wrapper feature selection techniques.

Table 2: Comparison of wrapper FS techniques in IDS.

Ref.	FS techniques	Datasets	Accuracy	Weakness
[38]	· GA · ELM	- NSL-KDD	87.53%	Only one dataset. Parameter Tuning. No Guarantee of Finding the Optimal Solution.
[39]	· GA · ELM · SVM	- IoT_ToN - UNSWNB15	99%	Computational complexity. Risk of overfitting. Parameter tuning.
[40]	· GA · GWO	- NSL-KDD	99.44%	Only one dataset. Stochastic Nature.
[41]	· PSO	- NSL-KDD	89.12%	Only one dataset. Parameter Tuning. Computational Complexity and Efficiency.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

[42]	· Sigmoid_PIO · Cosine_PIO	- NLS-KDD - UNSW-NB15	96%	Fitness and wights in fitness. Computational cost.
[43]	Bee colony technique utilizing Binary search equations Neural networks	- NSL-KDD	99%	Potential scalability issues. Risk of overfitting. Dependence on the quality of validation datasets.
[44]	· IG · CS · PSO	UNSW-NB15 Kyoto	99.25%	Balancing of dataset. Risk of Overfitting
[45]	· Decision tree · Min-Max normalization	- NLS-KDD - UNSW-NB15	99.87%	Bias Towards Features with More Levels. Instability
[46]	· GI · GA	- CIC-IDS2017 - CSE-CIC-IDS2018 - CIC-DDoS2019	99.98%	Modest accuracy improvement. GA can be computationally expensive.
[47]	· SVM · Random Forest Naïve Bayes	UNSW-NB15	93%	Computationally heavy. Tested on a single dataset.

4. Discussion

Two dominant families of approaches have shaped the development of feature selection in IDS research: filter-based methods and wrapper-based methods. Filter techniques rely on statistical or correlation measures to rank features independently of any classifier, offering speed and scalability in processing high-dimensional data. They are particularly attractive when real-time or resource-constrained detection is required, yet they may sacrifice accuracy by overlooking nonlinear interactions or feature synergies that are critical for detecting subtle patterns of intrusion. Wrapper methods on the other hand, utilize feature selection as part of a specific learning construction process, consider and evaluate different subsets of attributes on predictive performance in an iterative fashion. While the marriage of both these data sources is more effective as it takes relevance of individual features into account and in relation to other features but this incurs a huge computational burden. This trade-off between filters and wrappers highlight one of the main dichotomies in IDS research: efficiency versus accuracy. Recent approaches aim to reconcile this trade-off by combining filters and wrappers in hybrid strategies, which, at first order, apply filter to rapidly eliminate irrelevant features and then pass on the reduced subset of signal to a wrapper subsequently training classifiers with deeply optimized hyper-parameters. Different filter-based and wrapper-based methods have been presented for intrusion detection, in which no research work has revaluated the strengths of both together, any one at a time; quality wise as well as performance wise by comparing them with each other or against signature-based approach. Grounding it in the larger context of scalability, robustness and adaptability, it attempts to shed light on its strengths as well as the opportunities for work on IDS feature selection going forwards.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

Filter-based feature selection methods have been proven to be useful in IDS research for their capability of handling high-dimensional data efficiently and low computational cost. Chi-square, ANOVA, Information gain are good stats-based techniques used to check feature importance very often works really well and could show some of the state-of-the-art accuracies when benchmarked against these techniques on some standard data sets. However, this cannot be achieved and their major weakness comes when we compare with correlation-based methods. While ANOVA and information gain might miss some nonlinear dependencies and interactions, correlation-driven filters utilizing autoencoders or neural networks can uncover deeper insights on feature relevance that would either improve the detection-performance of an ML detector with decrease false alarm rates. However, even these correlation-based methods are sensitive to assumptions of linearity and the choices of thresholds. As they can eliminate biases to categorical variables or noise, fusion-based approaches are better than single-metric filters, but it is at the price of more costs on computation. In essence, thus, considered hybrid filters driven by correlations have a more general and robust base in comparison with the purely statistical ones that can still be attractive for the simplicity and speed.

The wrapper-based methods, on the other hand, adapt for accuracy and they incorporate feedback from the classifier into the selection process itself. Methods based on genetic algorithms, particle swarm optimization or grey wolf optimization always have an edge over filters in detection rates as they consider both the non-linear interaction of features and the search strategies within their optimization algorithms. For instances, GA with combination of ELM and SVM produces significantly superior accuracy / false positives when compared to statistical filters; yet, swarm-based methods like PSO allows for achieving comparable performance at lower feature counts. Workflows that incorporate wrappers and labelled-step 2: wrappers (e.g. GA and Gini impurity filter) also show huge dimensionality reduction capacity while also achieving high levels of predictive precision, which reflects the potential of specialised feature subsets in tailor-making feature subsets to a particular classifier in this case. However, the computational overhead introduced by wrappers is significant when being evaluated over large-scale or real-time IDS environments and hence less suitable for deployment scenario where speed and resource limitations are of importance.

Taken together, these two categories illustrate their complementarity. While filter methods are computationally efficient, they unable to learn some of the complex dependencies that are important to detect sophisticated intrusions. While wrapper methods are capable enough, it will be a cumbersome, time-consuming process, and they tend to overfit. Hybrid strategies that combine both categories represent a middle path where a filter stage first eliminates irrelevant or redundant features, reducing the search space and then a wrapper refines the subset further for classifier-specific optimization. These were tested in conjunction with existing pipelines, which use correlation-based filtering, and evolutionary or swarm-based wrappers, where the joint pipeline significantly outperformed either by itself. For example, pre-selecting features with correlation filters before applying GA or PSO lowers the cost of computation in a striking manner, while the accuracy improvements remain in line with wrapper methods. At the same time, a hybrid approach with statistical filters (e.g. information gain) together with optimization-based wrappers strikes an acceptable trade-off between running time of initial scoring and final iterative evaluation strength. Table 3 shows the Challenges and Proposed Solutions in Feature Selection Techniques for Intrusion Detection Systems.

These results demonstrate multiple directions for future work in the field of feature selection for intrusion detection. One initial avenue is joint usage of two distinguished pipeline — dotting filter and wrapper methods such that the use of both of them does not deteriorate accuracy but increases efficiency. This integration has already shown potential in hybrid models, though most of current designs do not consider adaptive mechanism either on the filter or wrapper side for dynamic adaptation. Integration of filter metrics and wrapper algorithms with the detection environment could inform a more adaptable framework, one that mirrors torsional forces on the wing so that as traffic patterns change and new attack lenses emerge over time, so too does its ability to detect anomalies. One major challenge is that we can do better than correlation-based filters. Though they are often successful at reducing dimensionality, these approaches still adhere to linear assumptions. Future research should investigate nonlinear correlation measures or deep learning-based embeddings with increased capability to model tangled relationships across features. Moreover, the efficacy of statistical filters like ANOVA and information gain can be improved by using them along with noise-robust mechanisms and interaction-aware scoring functions to decrease biases in heterogeneous data sources — improving transferability**.**

On the wrapper side, the challenge of computational cost remains central. Researchers should investigate lightweight optimization techniques or approximate search strategies that retain the accuracy benefits of evolutionary and swarm-based wrappers without incurring prohibitive overhead. Leveraging parallelization, distributed computing, and GPU acceleration could make wrappers more feasible in real-time IDS environments. In addition, integrating meta-learning

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

techniques to automatically tune the hyperparameters of optimization algorithms may reduce the risk of overfitting and improve the adaptability of wrapper-based feature selection.

Table 3: Challenges and Proposed Solutions in Feature Selection Techniques for IDS.

Feature Selection Technique	Type	Challenges	Proposed Solutions
Information Gain / ANOVA / Chi-Square	Filter	Bias towards features with many categories; sensitivity to noise and outliers; assumes linear relationships; ignores feature interactions	Combine with noise-robust measures; incorporate interaction-aware scoring; use fusion with other filters to reduce bias
Correlation-Based (CFS, CCFS)	Filter	Captures only linear dependencies; threshold sensitivity; affected by outliers; preprocessing limits (e.g., categorical encoding)	Adopt nonlinear correlation measures; robust thresholding; advanced encoding (e.g., one-hot, embeddings); hybrid with deep learning
Fusion-Based Filters (multi-metric approaches)	Filter	High computational cost; risk of overfitting in imbalanced data; limited real-time applicability	Develop lightweight fusion methods; integrate incremental/online updates; balance metrics adaptively
Genetic Algorithm (GA)	Wrapper	Computationally expensive; risk of overfitting; parameter tuning required; not guaranteed optimal solution	Use parallelization/GPU acceleration; integrate meta-learning for parameter tuning; hybrid with filters to shrink search space first
Particle Swarm Optimization (PSO)	Wrapper	Dataset dependency; limited exploration with fixed feature counts; still computationally demanding	Dynamic feature subset adaptation; cross-dataset validation; combine with pre-filtering for efficiency
Grey Wolf Optimizer (GWO)	Wrapper	Stochastic behavior; reliance on single datasets; high complexity	Hybridization with GA/PSO; validation across diverse datasets; design lightweight versions
Correlation + ANN (Hybrid)	Hybrid	Focuses on linear correlations only; computational complexity increases with ANN; threshold bias	Incorporate nonlinear correlation measures; use dimensionality reduction before ANN; adaptive threshold setting
Gini Impurity + GA (GIGA)	Hybrid	GA is computationally expensive; modest accuracy improvements over baseline	Enhance with distributed computing; explore alternative fitness functions; combine with faster evolutionary methods
Filter + Wrapper Hybrid Pipelines	Hybrid	Integration complexity; risk of redundancy; scalability to large-scale IDS	Adaptive integration frameworks; online incremental feature selection; automation with meta-learning to adjust balance dynamically

5. Conclusion and Future Work

Since feature selection is a crucial trade-off for any practical IDS that affects both the detection accuracy and computational efficiency, it is always one of the design principles behind an effective intrusion detection system.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

Through our survey, we have indicated that filter-based methods are scalable and swift which deems them practical for high-dimensional datasets, but they suffer from reduced capability in capturing sophisticated feature dependencies. Whereas wrapper-based methods outperform feature subset optimization for a specified classifier with higher accuracy levels, they are iterative in nature and hence, have substantial computational overhead. The following wide variety of approaches has sparked to one of the most widespread routes, hybrid strategies, where high dimensional data is filtered and this filter outputs are in turn optimized on by a wrapper resulting in performance-efficiency trade off.

Yet after decades of this growth, some perilous challenges still await. Hybrid models need adaptive mechanisms that can choose the combination of filter and wrapper based on traffic and attack pattern changes. Second, the limitations of correlation-based filters highlight the need for more powerful nonlinear dependence metrics or deep learning embedding is that had better capture mutual information across features. Processor scalability and real-time implementation are second major bottlenecks especially when the proposed wrapper method is applied at large-scale IDS environment. Hence, future work could aim to build light adaptive frameworks that can be auto-updated with statistical metrics using the power of statistical regression, MSA at both the local model labels as well as global label settings. Furthermore, integrating explainable AI to feature selection could increase the interpretability; hence, facilitate the creation of reliable and efficient intrusion detection systems.

Reference

[1] M. Du, "Application of information communication network security management and control based on big data technology," *Int. J. Commun. Syst.*, vol. 35, no. 5, p. e4643, 2022.

[2] X. Ma, X. Fu, B. Luo, X. Du, and M. Guizani, "A design of firewall based on feedback of intrusion detection system in cloud environment," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1–6.

[3] J. M. Kizza, "System intrusion detection and prevention," in *Guide to Computer Network Security*, 6th ed. Springer, 2024, pp. 295–323.

[4] N. A. Azeez, T. M. Bada, S. Misra, A. Adewumi, C. Van der Vyver, and R. Ahuja, "Intrusion detection and prevention systems: An updated review," in *Data Management, Analytics and Innovation* (Lecture Notes in Networks and Systems), vol. 1. Springer, 2020, pp. 685–696.

[5] E. F. E. Ahmet and I. N. Abaci, "Comparison of the host based intrusion detection systems and network based intrusion detection systems," *Celal Bayar Univ. J. Sci.*, vol. 18, no. 1, pp. 23–32, 2022.

[6] Y. Otoum and A. Nayak, "AS-IDS: Anomaly and signature based IDS for the internet of things," *J. Netw. Syst. Manag.*, vol. 29, no. 3, p. 23, 2021.

[7] M. Schrötter, A. Niemann, and B. Schnor, "A comparison of neural-network-based intrusion detection against signature-based detection in IoT networks," *Information*, vol. 15, no. 3, p. 164, 2024.

[8] C.-W. Chen, Y.-H. Tsai, F.-R. Chang, and W.-C. Lin, "Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results," *Expert Syst.*, vol. 37, no. 5, p. e12553, 2020.

[9] N. Kumar and U. Kumar, "Artificial intelligence for classification and regression tree based feature selection method for network intrusion detection system in various telecommunication technologies," *Comput. Intell.*, vol. 40, no. 1, p. e12500, 2024.

[10] M. Di Mauro, G. Galatro, G. Fortino, and A. Liotta, "Supervised feature selection techniques in network intrusion detection: A critical review," *Eng. Appl. Artif. Intell.*, vol. 101, p. 104216, 2021.

[11] F. Rahmat, A. Ali, S. Khan, M. Alazab, and T. R. Gadekallu, "Supervised feature selection using principal component analysis," *Knowl. Inf. Syst.*, vol. 66, no. 3, pp. 1955–1995, 2024.

[12] S. Bashir, A. S. Khan, Z. S. Khan, and F. A. Khan, "A novel feature selection method for classification of medical data using filters, wrappers, and embedded approaches," *Complexity*, vol. 2022, p. 9962023, 2022.

[13] P. Zhu, W. Zuo, L. Zhang, Q. Hu, and S. C. K. Shiu, "Unsupervised feature selection by regularized self-representation," *Pattern Recognit.*, vol. 48, no. 2, pp. 438–446, 2015.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

[14] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "A review of recent approaches on wrapper feature selection for intrusion detection," *Expert Syst. Appl.*, vol. 198, p. 116822, 2022.

[15] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, M. Vijayalakshmi, P. Yogesh, and A. Kannan, "Intelligent feature selection and classification techniques for intrusion detection in networks: A survey," *EURASIP J. Wirel. Commun. Netw*, vol. 2013, no. 1, p. 271, 2013.

[16] J. Maldonado, M. C. Riff, and B. Neveu, "A review of recent approaches on wrapper feature selection for intrusion detection," *Expert Syst. Appl.*, vol. 198, p. 116822, 2022.

[17] H. K. Malik, N. J. Al-Anber, and F. A. E. Al-Mekhlafi, "Comparison of feature selection and feature extraction role in dimensionality reduction of big data," *J. Tech.*, vol. 5, no. 1, pp. 1–8, 2023.

[18] B. K. Padhi, S. Chakravarty, B. Naik, R. M. Pattanayak, and H. Das, "RHSOFS: Feature selection using the rock hyrax swarm optimization algorithm for credit card fraud detection system," *Sensors*, vol. 22, no. 23, p. 9321, 2022.

[19] I. B. Adnan, M. A. Khan, and R. A. Khan, "A novel approach for anomaly detection in IoT using hybrid feature selection techniques," *IEEE Access*, vol. 12, pp. 12345–12358, 2024.

[20] G. Ansari, T. Ahmad, and M. N. Doja, "Hybrid filter–wrapper feature selection method for sentiment classification," *Arab. J. Sci. Eng.*, vol. 44, no. 10, pp. 9191–9208, 2019.

[21] J. Abawajy, A. Darem, and A. A. Alhashmi, "Feature subset selection for malware detection in smart IoT platforms," *Sensors*, vol. 21, no. 4, p. 1374, 2021.

[22] Y. B. Wah, N. Ibrahim, H. A. Hamid, S. Abdul-Rahman, and S. Fong, "Feature selection methods: Case of filter and wrapper approaches for maximising classification accuracy," *Pertanika J. Sci. Technol.*, vol. 26, no. 1, pp. 329–340, 2018.

[23] V. Herrera-Semenets, L. Bustio-Martínez, R. Hernández-León, and J. Van den Berg, "A multi-measure feature selection algorithm for efficacious intrusion detection," *Knowl.-Based Syst.*, vol. 227, p. 107264, 2021.

[24] S. Alhassan, G. Abdul-Salaam, A. Micheal, Y. M. Missah, E. D. Ganaa, and A. S. Shirazu, "CFS-AE: Correlation-based feature selection and autoencoder for improved intrusion detection system performance," *Unpublished*.

[25] G. Farahani, "Feature selection based on cross-correlation for the intrusion detection system," *Secur. Commun. Networks*, vol. 2020, p. 8844661, 2020.

[26] I. S. Thaseen, J. S. Banu, K. Lavanya, M. R. Ghalib, and K. Abhishek, "An integrated intrusion detection system using correlation-based attribute selection and artificial neural network," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 2, p. e4014, 2021.

[27] R. Zhao, Y. Mu, L. Zou, and X. Wen, "A hybrid intrusion detection system based on feature selection and weighted stacking classifier," *IEEE Access*, vol. 10, pp. 71414–71426, 2022.

[28] S. Venkatesan, "Design an intrusion detection system based on feature selection using ML algorithms," *Math. Stat. Eng. Appl.*, vol. 72, no. 1, pp. 702–710, 2023.

[29] M. J. Siraj, T. Ahmad, and R. M. Ijtihadie, "Analyzing ANOVA F-test and sequential feature selection for intrusion detection systems," *Int. J. Adv. Soft Comput. Its Appl.*, vol. 14, no. 2, pp. 1–15, 2022.

[30] A. Thakkar and R. Lohiya, "Fusion of statistical importance for feature selection in deep neural network-based intrusion detection system," *Inf. Fusion*, vol. 90, pp. 353–363, 2023.

[31] P. Nimbalkar and D. Kshirsagar, "Feature selection for intrusion detection system in Internet-of-Things (IoT)," *ICT Express*, vol. 7, no. 2, pp. 177–181, 2021.

[32] M. A. Siddiqi and W. Pak, "Optimizing filter-based feature selection method flow for intrusion detection system," *Electronics*, vol. 9, no. 12, p. 2114, 2020.

[33] I. Karna, A. Madam, C. Deokule, R. Adhao, and V. Pachghare, "Ensemble-based filter feature selection technique for building flow-based IDS," in *Proc. 2nd Int. Conf. Adv. Comput., Commun., Embedded Secure Syst. (ACCESS)*, Mumbai, India, 2021, pp. 324–328.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025

[34] N. El Aboudi and L. Benhlima, "Review on wrapper feature selection approaches," in *Proc. Int. Conf. Eng. MIS (ICEMIS)*, Agadir, Morocco, 2016, pp. 1–5.

[35] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," *Appl. Soft Comput.*, vol. 62, pp. 441–453, 2018.

[36] B. Nouri-Moghaddam, M. Ghazanfari, and M. Fathian, "A novel multi-objective forest optimization algorithm for wrapper feature selection," *Expert Syst. Appl.*, vol. 175, p. 114737, 2021.

[37] W. L. Al-Yaseen, A. K. Idrees, and F. H. Almasoudy, "Wrapper feature selection method based differential evolution and extreme learning machine for intrusion detection system," *Pattern Recognit.*, vol. 132, p. 108912, 2022.

[38] F. H. Almasoudy, W. L. Al-Yaseen, and A. K. Idrees, "Differential evolution wrapper feature selection for intrusion detection system," *Procedia Comput. Sci.*, vol. 167, pp. 1230–1239, 2020.

[39] E. M. Maseno and Z. Wang, "Hybrid wrapper feature selection method based on genetic algorithm and extreme learning machine for intrusion detection," *J. Big Data*, vol. 11, no. 1, p. 24, 2024.

[40] N. Kunhare, R. Tiwari, and J. Dhar, "Intrusion detection system using hybrid classifiers with meta-heuristic algorithms for the optimization and feature selection by genetic algorithm," *Comput. Electr. Eng.*, vol. 103, p. 108383, 2022.

[41] N. Kunhare, R. Tiwari, and J. Dhar, "Particle swarm optimization and feature selection for intrusion detection system," *Sādhanā*, vol. 45, no. 1, p. 109, 2020.

[42] H. Alazzam, A. Sharieh, and K. E. Sabri, "A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer," *Expert Syst. Appl.*, vol. 148, p. 113249, 2020.

[43] I. Laassar, M. Y. Hadi, H. R. Arifullah, and F. S. Khan, "Proposed algorithm base optimisation plan for feature selection-based intrusion detection in cloud computing," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 2, pp. 1140–1149, 2024.

[44] M. Bakro, A. Al-Sarem, A. Saeed, F. Saeed, W. Elmedany, and M. Alazab, "An improved design for a cloud intrusion detection system using hybrid features selection approach with ML classifier," *IEEE Access*, vol. 11, pp. 64228–64247, 2023.

[45] M. A. Umar, Z. Chen, K. Shuaib, and Y. Liu, "Effects of feature selection and normalization on network intrusion detection," *Authorea Prepr.*, 2024.

[46] R. Lalduhsaka and A. K. Khan, "Enhanced intrusion detection system using a two-staged feature selection method," *Secur. Priv.*, vol. 8, no. 3, p. e70025, 2025.

[47] M. Umer, M. Tahir, M. Sardaraz, M. Sharif, H. Elmannai, and A. D. Algarni, "Network intrusion detection model using wrapper based feature selection and multi head attention transformers," *Sci. Rep.*, vol. 15, no. 1, p. 28718, 2025.

DOI: https://doi.org/10.54216/JCIM.170208
Received: April 04, 2025 Revised: June 17, 2025 Accepted: August 12, 2025