Received: 6 October 2024 | Revised: 14 January 2025 | Accepted: 13 February 2025 | Published online: 28 March 2025

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2026, Vol. 5(1) 143–151
DOI: 10.47852/bonviewJCCE52024492

BON VIEW PUBLISHING

A New Hybrid Approach for Improving Intrusion Detection System Based on Scatter Search Algorithm and Support Vector Machine

Feras E. AbuAladas[1], Mohammad Shehab[1,*], Rasha Israwah[1], Ghaith Jaradat[1] and Yousef Qawqzeh[2]

[1] *College of Information Technology, Amman Arab University, Jordan*

[2] *Information Technology College, University of Fujairah, United Arab Emirates*

Abstract: In recent years, the extensive growth of the internet, coupled with the integration of sensors and wireless sensor networks into critical areas like healthcare and military defense, has led to a significant expansion in the use of artificial intelligence and the Internet of Things applications. Due to the importance and sensitivity of the data in these fields, it is crucial to use an intrusion detection system (IDS), which applies specific algorithms to analyze and process the data from the network in order to find any suspicious activities or behaviors on the system and improve system security. A network frequently carries enormous amounts of data, particularly in military applications where data must flow continuously. In this research, an enhanced IDS using the scatter search algorithm is proposed. A population of randomly generated initial solutions is used to produce a diverse group of selected and high-performing solutions, which serve as a reference set to steer the search process. This reference set is then adopted as a feature selection method based on a support vector machine. The efficiency of the suggested method was examined utilizing the NSL-KDD dataset, and the results were compared with the Gazelle Optimization Algorithm, Algorithmic Optimization Algorithm, Gray Wolf Optimizer, Adjusted Gray Wolf Optimizer, and Particle Swarm Optimization. The main performance metrics utilized to test the efficiency of the suggested method include accuracy, detection efficiency, false-positive rate, and feature count. The results illustrated that the suggested method has obtained a high intrusion detection accuracy in the IDS system of 99% and decreased false alarm rates (0.02) and selected only 17 features from the initial dataset, which contained 41 features, demonstrating the effectiveness of the suggested method.

Keywords: scattered search algorithm, intrusion detection system, Internet of Things, feature selection, support vector machine, optimization algorithm

1. Introduction

The widespread use of the Internet and its integration into critical domains such as healthcare, military, and financial services have made the protection of data and systems from malicious attacks a top priority [1]. The accelerated development of technologies such as artificial intelligence, Internet of Things (IoT), and cloud computing has also contributed to heightened exposure to security vulnerabilities [2]. As networks handle increasingly large amounts of data, intrusion detection systems (IDSs) are essential for protecting the pillars (confidentiality, integrity, and availability) of these systems. The IDS is a security tool that monitors and identifies suspicious or unauthorized actions within a network. It can be classified into two major groups depending on the technique of detection: signature-based and anomaly-based systems. Signature-based IDS functions by recognizing established threat patterns, whereas anomaly-based IDS utilizes machine learning (ML) techniques to detect abnormal activity, enabling them to uncover novel threats [3].

Despite their benefits, conventional IDS encounter multiple challenges, such as elevated false-positive rates, significant computational demands, and difficulty identifying sophisticated attacks. Furthermore, as cyberattacks grow more complex, including Distributed Denial of Service (DDoS), zero-day exploits, and Advanced Persistent Threats, there is an increasing demand for more advanced and flexible IDS [4].

Recent developments in ML have introduced a novel approach to creating more effective IDS models. The effectiveness of IDS can be greatly improved through the use of optimization algorithms and feature selection methods [5]. Feature selection plays a crucial role in simplifying the dataset by identifying and retaining only the most important features, which enhances detection accuracy and reduces the likelihood of false positives.

This research presents a progressing IDS by coordinating the scatter search (SS) algorithm [6] with a support vector machine (SVM). SS could be a meta-heuristic approach that works on a variety of arrangements and is recognized for its capacity to address challenging optimization assignments proficiently. This work utilized SS for feature selection to pinpoint the key qualities from the dataset, which are then classified utilizing SVM. This

***Corresponding author:** Mohammad Shehab, College of Information Technology, Amman Arab University, Jordan. Email: m.shehab@aau.edu.jo

integration permits our framework to improve location precision, minimize wrong positives, and diminish computational requests.

The NSL-KDD dataset is utilized for assessing the viability of interruption location frameworks [7]. The assessment of the SVM-SS demonstrates its performance against several leading optimization techniques, including the Gazelle Optimization Algorithm (GOA), Algorithmic Optimization Algorithm (AOA), Gray Wolf Optimizer (GWO), Adjusted Gray Wolf Optimizer (AGWO), and Particle Swarm Optimization (PSO). The results show that the SVM-SS outperforms these methods in terms of accuracy, detection rate, distance, and resolution, highlighting its potential as an effective solution for practical intrusion detection applications.

2. Literature Review

Recently, various approaches have been implemented in multiple IDS models to enhance their security. For example, researchers have developed a range of meta-heuristic techniques, including the Remora Optimization Algorithm, aimed at improving IDS performance through feature selection and boosting detection accuracy [8]. In addition, there are many works adapted to ML to solve various problems [9, 10]. Various ML approaches can be employed to develop anomaly-based IDS. The two primary approaches frequently employed are oversee learning and non-oversee learning. In oversee learning, a mapping function is employed to align designated input–output pairs. Non-oversee learning enables a model to autonomously identify inherent relationships within the data. In contrast, SVM, which is a form of oversee ML, is commonly employed in IDS for the classification of attacks. K-nearest neighbors (KNN) and decision tree (DT) are more instances of those ML techniques. On the other hand, clustering algorithms are usually directed by unsupervised ML techniques like K-means [11].

The role of shallow ML models, such as DTs, KNN, and SVM, in addressing computational efficiency in low-resource environments is summarized in the following key studies, including the optimization of microfluidic synthesis of silver nanoparticles [12], which highlights how optimization techniques like SS can be applied in contexts similar to IDS. The study on acoustic emission and ML algorithms for particle size analysis [13] highlights the importance of feature selection and shallow learning in improving prediction accuracy. Parameter flexible wildfire prediction using ML techniques [14], which showcases the power of ML-based feature selection and anomaly detection.

Network administrators and researchers face challenges in ensuring networked computer security. Studies have proposed a framework combining ML techniques with attribute selection algorithms for effective intrusion detection. The model uses a hybrid meta-heuristic feature selection technique and supervised ML algorithms for enhanced detection accuracy, execution speed, and error rate utilizing the NSL-KDD dataset.

In Türk [15], the authors used the UNSW-NB15 and NSL-KDD datasets for a thorough attack detection process using current ML methods. Two-class and multi-class accuracy in the UNSW-NB115 dataset were, respectively, 98.6% and 98.3%, while precision in the NSL-KDD dataset was 97.8% and 93.4%. The findings demonstrate the significance of ML techniques in enhancing IDS.

Gaye et al. [16] proposed the ML classification algorithm (SVM) as one of the machine-learning-based IDS. In order to help the classifier handle fewer support vectors and reduce structural risk, it is motivated by approaching the dual formulation of high-dimensional challenges.

An IDS technique called lightweight Intelligent Intrusion Detection Model for Wireless Sensor Network (WSN) was introduced by Pan et al. [17]. The suggested approach integrates the sine cosine algorithm (SCA) and the kNN, which significantly improved detection accuracy and reduced the false alarm rate (FAR), resulting in an effective IDS. In order to speed up processing, the chaos sine and cosine algorithm (CSCA) was also given a compact mechanism. On the other hand, accuracy is improved using the polymorphic mutation approach. Using the NSL-KDD and UNSW-NB15 datasets for testing, the constructed IDS model produced very effective results.

Akhtar and Feng [18] worked with the NSL-KDD dataset; they produced a convolutional neural network (CNN)-based DoS security breach detection mechanism that performed well in DoS. However, they could only identify known DoS network violations. Also, using prior research as a guide, we came to the conclusion that using optimization techniques to pick out the most pertinent features would increase the accuracy of the IDS model, decrease its computational complexity, and lessen over- and under-fitting problems [19].

A method for feature selection in IDSs to detect DoS and DDoS offensives was proposed in Nimbalkar and Kshirsagar [20]. The proposed system employs information gain (IG) and gain ratio (GR) to prioritize the top half of the features; this method outperforms traditional IDSs using fewer features on the IoT-BoT and KDD Cup 1999 datasets.

In the optimal SVM (OSVM) model presented by Amaran and Mohan [21], there are three subprocesses in the proposed model: preprocessing, classification, and kernel selection for the IDS in WSNs. First, after preprocessing the input data in the network so that it is in a convenient format, in the second subprocess, the optimal kernels in the SVM are selected proficiently by the whale optimization algorithm (WOA). Then, finally, the OSVM model to classify the intrusions in the NSL-KDD CUP 99 dataset was applied to the experiments. The results were 94.09% for accuracy and 95.02% for detection rate.

Nugroho et al. [22] proposed an IDS based on Spark and Conv-AE that makes use of open datasets like KDD 99. The results show that unbalanced datasets impair model performance. A method for selecting relevant features called dynamic recursive feature selection algorithm (DRFSA) was created by Nancy et al. [23] to determine the ideal number of features for the IDS process. Additionally, fuzzy temporal constraints were added to a modified decision tree method to enhance the precision of the classification of network data. Furthermore, CNNs were utilized for the classification of big data. KDD Cups were used in the method's experimentation.

The developed model has a higher intrusion detection rate, a better packet delivery ratio, and a lower throughput, according to the results. Additionally, the FAR is lower.

According to established research, the bulk of IDS optimization strategies in the literature attempt to solve the IDS's poor accuracy problem. In this research, we developed a way for choosing the most pertinent features using the SS algorithm, and we found that it performed better than other approaches and techniques in the literature in terms of accuracy, FAR, feature count, and detection rate. Table 1 presents a comparison of various techniques used in IDS.

Existing IDSs face high false-positive rates, inefficiency, and challenges detecting complex attacks like zero-day exploits. Traditional optimization methods often fail with high-dimensional datasets. The proposed hybrid approach combines the SS Algorithm and SVM, reducing features to 17 and achieving 99% accuracy with a 0.02 FAR. The feature selection process leads to a more robust and lightweight IDS, making it applicable to real-time and large-scale

Table 1
Comparison of different techniques for ID

Ref	Dataset	Methodology	Results
[15]	NSL-KDD	Evaluation of IDS Using ML on UNSW-NB15 and NSL-KDD Datasets	Accuracy 97.8% and 93.4%
[16]	Big data	Enhancing SVM Algorithms in the Context of Big Data	Accuracy 98%
[17]	NSL-KDD and UNSW-NB15	A Lightweight Intelligent Intrusion Detection Model for Wireless Sensor Networks	The model can be used for a cloud computing structure and fog computing, energy consumption, and FAR
[18]	NSL-KDD	CNN-based DoS intrusion detection model was produced from the dataset, and it performed well in DoS but could only identify known DoS network assaults	Accuracy of 98% and 99%
[20]	KDD Cup 1999	Selecting Features for IDS on the Internet of Things (IoT)	A feature selection method utilizing IG and GR outperforms conventional IDSs in identifying DoS and DDoS
[21]	NSL-KDD Cup 99	Optimized SVM-Based IDS for Wireless Sensor Networks	The OSVM model employs an efficient method for choosing the best kernels in the SVM framework using WOA to detect intrusions
[24]	NSL-KDD dataset	CNN-based DoS intrusion detection model was produced from the dataset, and it performed well in DoS but could only identify known DoS network assaults	Accuracy of 98% and 99%
[22]	KDD Cup 99, NSL-KDD	Identifies the types or categories of attacks that the IDS uses to evaluate whether an intrusion has occurred	ANN 16%, RNN 12%
[23]	KDD Cup	Method for feature selection called DRFSA	The developed model has a more accurate intrusion detection rate, better packet delivery ratio, and less throughput

datasets. Real-world attack scenarios, such as zero-day attacks, are provided as cases where SVM-SS could improve detection.

3. Proposed Method

Most of the traditional IDS models have good but not satisfactory performance measures regarding accuracy and detection rate. This research introduces more advanced models to overcome such issues by using SS for the feature selection process. After merging it with SVM to enhance the local search process, this enhanced the performance of IDS, where the proposed approach focusing on employing SS to choose the best features of the set to be used by SVM for IDS is further explained as shown in Figures 1 and 2.

As shown in Figures 1 and 2, the step starts to prepare and normalize NSL-KDD by converting all string values into integer numbers. After that, the ideal set of features is selected via SVM-SS as presented by two main phases: the starts with the first phase, which is population initialization, and the second phase, which includes the exploration processes. Finally, the set of features selected by SS is used by the SVM to classify the input data into abnormal and normal, where the performance metrics (number of selected features, accuracy, FAR, and detection rate) calculated depend on the SVM results.

SS is a meta-heuristic optimization algorithm used for solving combinatorial optimization problems. The SS algorithm is a search strategy in which a set of optimizations within a cluster cooperate and compete. SS creates an initial population of random answers and methodically chooses a collection of elite, diversified solutions to serve as a reference set for the search. The efficiency of the IDS is improved by increasing the accuracy of classification or recognition, which is done through the use of the SS algorithm with the SVM to choose the most appropriate features in the data and address the imbalanced features or redundant correlations.

4. Experimental Results

4.1. Dataset

In this work, an updated NSL-KDD dataset of KDD Cup 99 proposed by Ghorbani et al. [25] is considered as benchmark to investigate the efficiency of the suggested method. The NSL-KDD dataset has forty-one (41) features, as shown in Tables 2, 3, and 4. These sets can be categorized as either normal or attack, where features are divided into three types based on their characteristics: nominal, binary, and numeric.

The NSL-KDD dataset has four main types of attacks:

Denial of Service Attacks (DoS): When a person is blocked from using a service because of a hostile action.

User-to-Root Attacks (U2R): When an unauthorized individual gains illicit access to the main computer system.

Remote to local attacks (R2L): An unauthorized user accesses the main system without permission.

Probing attacks: When an attacker scans the network to collect information about the system for the purpose of evading security protocols.

Figure 1
The proposed detection method

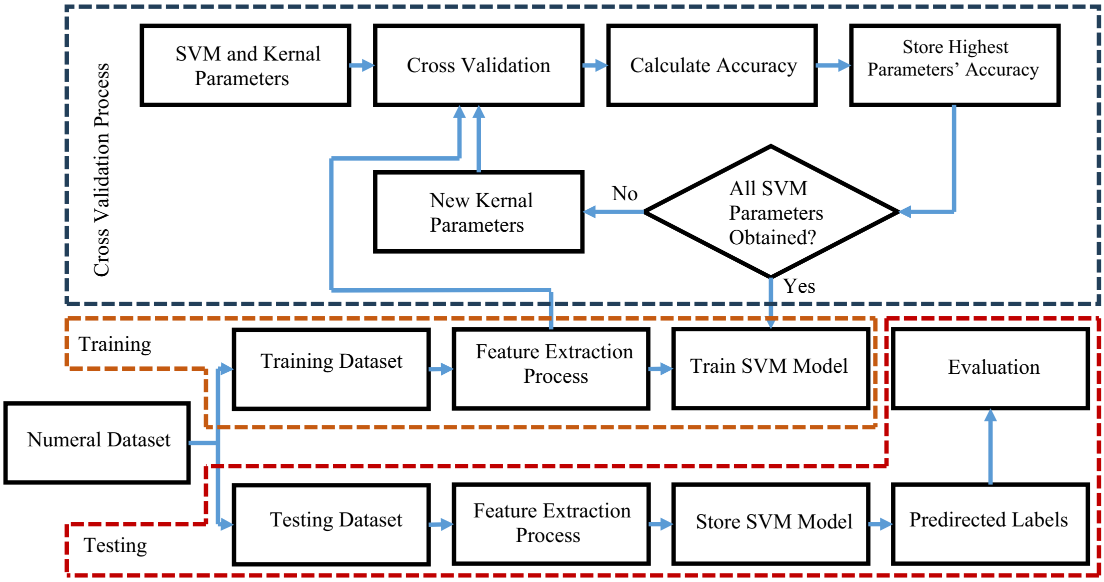


Figure 2
A flowchart of SVM training and evaluating process

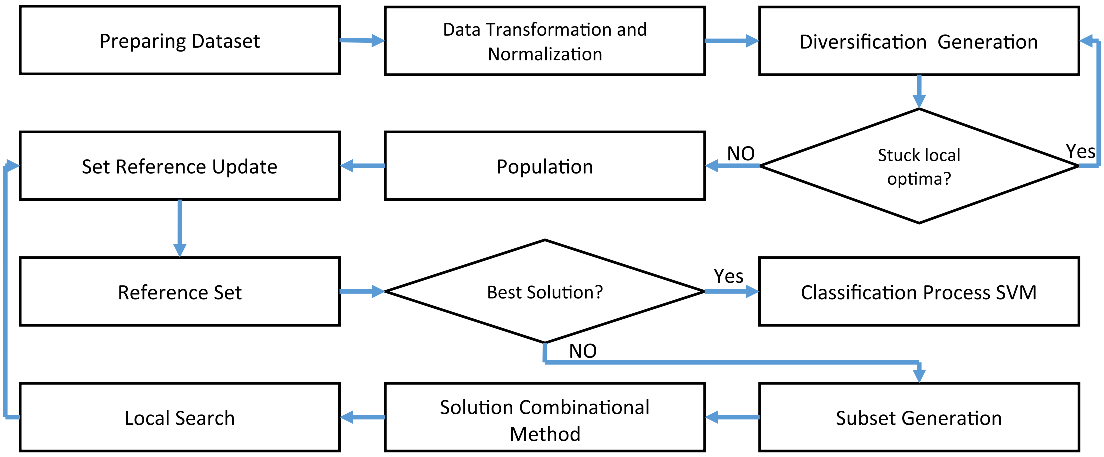


Table 2
Features of NSL-KDD dataset/basic category

Name	Data type
duration	constant
'protocol_type'	emblematic
service	emblematic
flag	emblematic
Scr_bytes	constant
Dst_ bytes	constant
land	emblematic
Wrong_fragment	constant
Urgent	constant

The NSL-KDD dataset was selected for method analyses because it contains less repetitive data compared to the KDD 99 dataset. This makes it more representative of real-world scenarios. The dataset consists of 125,973 network flow records, 41 features, and a single class indicating whether the record is an attack or normal. The advantages of using the NSL-KDD dataset include the removal of redundant records, which allows classifiers to yield unbiased results, and a sufficient number of records in the training and evaluation datasets, making it reasonable to perform experiments on the whole set. Additionally, from each difficulty level of the group, the proportion of records chosen decreases as the percentage of records in the original KDD dataset increases.

Before utilizing the dataset for testing, several steps must be taken. First, categorical features are converted to numeric values to make them readable by the SVM, as illustrated in Table 5, which shows the values of features alongside their corresponding numeric values utilized for this transformation. Next, normalization

Table 3
Features of NSL-KDD dataset/content category

Name	Data type
Hot	constant
Num_failed_logins	constant
Logged_in	emblematic
Num_compromised	constant
Root_shell	constant
Su_attempted	constant
Num_root	constant
Is_guest_login	emblematic
count	constant
Srv_count	constant
Serror_rate	constant
Srv_serror_rate	constant
Rerror_rat	constant
Srv_rerror_rate	constant
Same_srv_rate	constant
Diff_srv_rat	constant
Srv_ diff_srv_rat	constant
Num_file_creation	constant
Num_shells	constant
Num_access_files	constant
Num_outbound_cmd	constant
Is_host_login	emblematic

Table 4
Features of NSL-KDD dataset/traffic category

Name	Data type
Dst_host_count	constant
Dst_host_srv_rat	constant
Dst_host_same_srv_rate	constant
Dst_host_diff_srv_rat	constant
Dst_host_same_srv_port_rate	constant
Dst_host_srv_diff_host_rate	constant
Dst_host_serror_rat	constant
Dst_host_srv_serror_rat	constant

Table 5
Transform methodology

Attribute value	Attribute
tcp	‘protocol_type’
icmp	‘protocol_type’
udp	‘protocol_type’
REJ	‘flag’
SF	‘flag’
RSTO	‘flag’
S0	‘flag’
domain	‘service’
name	‘service’
pop_2	‘service’
http_443	‘service’
anomaly	‘class’

is applied to the data to standardize the domain of feature values between 0 and 1. Given the non-uniform distribution of features, the max-min normalization technique is applied to the NSL-KDD dataset, as described in the equation:

$$X' = (Original\ value\ Min\_Value)/(Maxvalue - Min\_Value) \quad (1)$$

This transformation was necessary to enable the SVM algorithm to process and interpret the data effectively. The table displays four categorical attributes: ‘protocol_type’, ‘flag’, ‘service’, and ‘class’. Each attribute value has been assigned a corresponding numerical representation. For example, in the ‘protocol_type’ attribute, ‘tcp’ is represented by 1, ‘udp’ by 3, and ‘icmp’ by 2. Similarly, in the ‘flag’ attribute, ‘REJ’ is assigned a value of 1, ‘SF’ is assigned 2, and so on. The ‘service’ attribute also follows the same pattern, with categorical values such as ‘private’ mapped to 1, ‘other’ mapped to 17, and the remaining values assigned subsequent numerical representations. Additionally, the table includes the ‘class’ attribute, which denotes the class labels of the dataset. In this attribute, the value ‘anomaly’ is represented by 1, indicating the presence of abnormal or malicious network activity. Conversely, the value ‘normal’ is assigned a numerical representation of 0, indicating the absence of any anomalous behavior.

This change encourages the SVM’s capacity to prepare the dataset because it requires numerical inputs. By giving a clear mapping between categorical trait values and their comparing numerical representations, the table improves the understandability and convenience of the NSL-KDD dataset in the context of intrusion detection analysis and research. In the next step, the SVM-SS method is used to identify the optimal feature set for classification using SVM.

The computational complexity of SVM can be computationally expensive, especially for large datasets, as it involves solving a convex optimization problem that requires solving a set of dual variables, where SVM performance can be sensitive to the choice of hyperparameters, such as the regularization parameter and the kernel parameters, which may require careful tuning to achieve optimal results.

4.2. Experiment setup

In this work, the effectiveness of the suggested approach is evaluated by experimenting using NSL-KDD as a test dataset, and the outcomes of the proposed approach are evaluated against those of other methods GOA [26], AOA [27], GWO [28], AGWO [29], and PSO [30] methods considering the evaluation metrics. The test simulation was carried out on a PC with a Core i7 2.4 gigahertz CPU, 8 GB RAM, and using PYTHON COLAB. The experimental results are evaluated using the following metrics as shown in Table 6.

$$Accuracy = \frac{(TP + TN)}{TP\ +\ TN\ +\ FP\ +\ FN} \ *\ 100\% \quad (2)$$

Detection Rate (DR): It is the rate of (TP) samples to the (TP + FN).

$$DR = \frac{(TP)}{(TP + FN)} \quad (3)$$

Table 6
Evaluation metrics

Evaluation metrics	Description
True Positive (TP)	Occurs when the system correctly identifies an actual intrusion
True Negatives (TN)	When the system correctly classifies typical behavior as non-intrusive
False Positives (FP)	When the system wrongly identifies normal behavior as an intrusion
False Negatives (FN)	When a system misses a real incursion
Accuracy	Refers to the percentage of correctly identified data true positives (TP) and true negatives (TN) compared to the overall number of samples (TP + TN + FP + FN)

False Alarm Rate (FAR): It is the rate of (FP) to the total of the total amount of non-intrusive samples (FP + TN).

$$FAR = \frac{(FP)}{(FP + TN)} \tag{4}$$

Number of Features: The number of features has been chosen for use in the classification process to classify the events into normal and abnormal activities.

4.3. Results and discussion

The study explores and compares the test results of the proposed technique with those of GOA, AOA, GWO, AGWO, and PSO in terms of performance indicators like precision, success rate of detection, feature count, and FAR. Each experiment was repeated 30 times to ensure statistical reliability, and the results were analyzed using ANOVA to confirm their statistical significance. The additional insights highlight the robustness of SVM-SS in achieving superior accuracy and feature reduction compared to existing methods, as shown in the extended results table and corresponding discussion. It's worth to mention that the parameters setting of the selected techniques are shown in Table 7.

Table 7
Parameters setting

Alg.	Parameter	Range
GOA AOA GWO PSO	f	[0.1]
	l	[0.00001, 1]
	α	[0.2, 1]
	μ	Dynamic
	α	[0, 2]
	A,C	[0, 2]
	c1,c2c_1, c_2c1,c2	2.0
	W	0.4–0.9
	V	0.2–0.8

Overall, the experimental results indicate that the suggested method, SVM-SS, has a significant positive impact on the feature selection process, as well as on the overall performance of the IDS. Its ability to select relevant features, combined with its high accuracy, detection rate, and low FAR, makes SVM-SS a promising approach for IDS as shown in Table 8.

Furthermore, the proposed method (SVM-SS) achieved the value of 0.02 in FAR analysis, as shown in Figure 3(a), which is equal to GOA, but less than AOA, AGWO, GWO, and PSO with 0.03, 0.03, 0.24, and 0.26, respectively, which indicates that SVM-SS and GOA have less error diagnosing normal data as suspicious.

The analysis shows that the SVM-SS method has a detection rate higher than the GWO method and close to that of the GOA algorithm and AOA with a slight difference. Its detection rate is slightly less than GOA, AOA, and AGWO with 98%, 97.9%, and 96.9% respectively. Nevertheless, the SVM-SS detection rate is higher than GWO, AGWO, and PSO, with values of 83%, 96%, and 93%, respectively. As a result, the SVM-SS is the best in terms of detection rate, and this results in much better IDS performance as shown in Figure 3(b). Figure 3(c) presents the number of features selected by SVM-SS and other algorithms. The SVM-SS method selected 18 features, while GWO achieved the highest number with 23 features, and AGWO selected the fewest with 12 features.

Scalability and real-time application of SVM-SS are highlighted in three main aspects. (i) Computational complexity: where the complexity of the SS algorithm increases with the dimensionality and size of the dataset. In larger datasets, this increase affects the execution time. To mitigate this, parallelized implementations of SS and SVM can be utilized to process large datasets. Techniques like parallel processing, GPU acceleration, and distributed computing using frameworks like Apache Spark can be proposed. (ii) Real-time detection: where real-time detection requires fast feature selection and classification. This challenge can be met by employing incremental learning strategies, where the model updates itself as new data arrives, rather than retraining from scratch. Real-time optimizations also involve reducing response latency by limiting the number of selected features. (iii) Practical scenarios such as intrusion detection in smart grids, autonomous vehicles, and critical infrastructures.

The generalization of SVM-SS beyond cybersecurity is pointed out as follows. A cross-domain application highlights how SVM-SS can be applied to domains like financial fraud detection, healthcare anomaly detection, and manufacturing fault detection. While feature selection is a universal need in ML-driven anomaly detection, SVM-SS can reduce feature dimensionality and improve classification accuracy in these domains. An example of fraud detection emphasizes that feature selection can reduce computational overhead while improving anomaly detection in large datasets like financial transaction logs. Thus, suggestions for future research include testing the SVM-SS on anomaly detection datasets from other fields, such as banking and e-commerce. While deep learning models like CNNs offer strong detection accuracy, SVM-SS excels in low-resource environments, is computationally cheaper, and is suitable for embedded and edge devices.

From Table 5, it can be inferred that the suggested SVM-SS method has a value of 99.00%, which exceeds the accuracy values of all compared methods, while GOA has the next higher value of

Table 8
Experimental results of the tested methods

Evaluation Metric	SVM-SS	GOA	AOA	AGWO	GWO	PSO
Accuracy	99%	98%	97.50%	96%	79%	89%
Detection rate	99%	97. 90%	96.90%	96%	83%	93%
False alarm rate	0.02	0.02	0.03	0.03	0.24	0.26
No. features	17	23	14	12	27	20

Figure 3
Accuracy values of compared methods

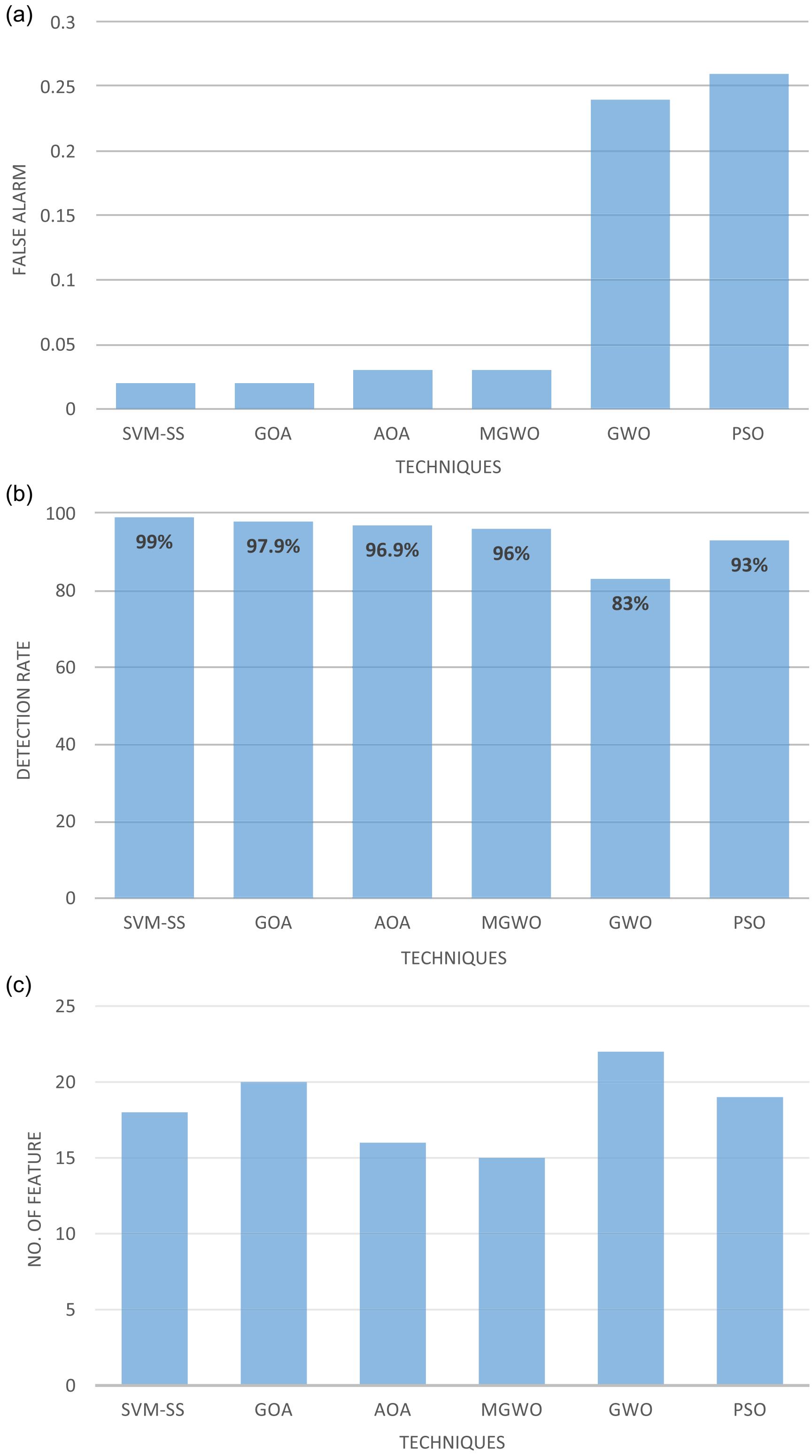
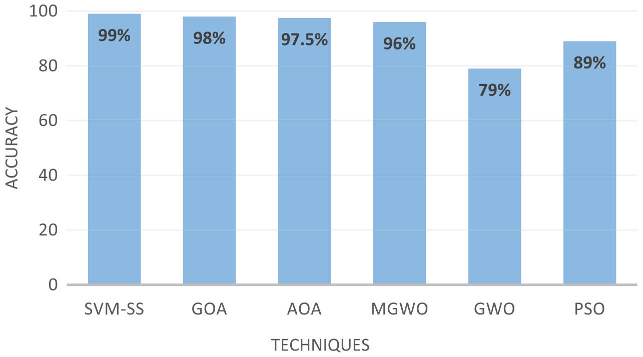


Figure 4
Accuracy values of compared methods



98.00% accuracy. GOA, AOA, AGWO, GWO, and PSO have accuracies of 97.5%, 96%, 79%, and 89%, respectively. This indicates the excellence of SVM-SS, in precision over the other compared methods, which reflects that the proposed SVM-SS method has the great ability to distinguish normal and suspicious data among the compared methods. Figure 4 shows the accuracy values of the compared methods.

5. Conclusion and Future Work

In this research, a method to improve the efficiency and precision of IDS was proposed using the SS algorithm for feature selection and the SVM classifier. The proposed SVM-SS approach's effectiveness is evaluated by experimentation using the test dataset NSL-KDD, and the outcomes are scored based on precision, identification success, FAR, and the quantity of chosen attributes. The proposed method's experimental results are compared with those of the GOA, AOA, GWO, AGWO, and PSO methods. It is found that SVM-SS is more effective than the rest of the compared methods regarding assessment criteria, particularly concerning precision, which suggests that the proposed method is much more effective as a feature selection tool and produces an IDS that is well-designed and results in a more secure network.

The proposed method, SVM-SS, had a significant impact on the feature selection process by enhancing its accuracy. Through the application of SVM-SS, 17 appropriate features were selected from a total of 41 features, which resulted in an excellent accuracy rate. This improvement was empirically validated, demonstrating the efficacy of SVM-SS in identifying the most relevant features for intrusion detection. Moreover, the application of SVM-SS had a profound effect on the comprehensive effectiveness of the IDS. The enhanced feature selection process contributed to the increased quality of the IDS by achieving a high detection rate and a low FAR. Consequently, the IDS successfully identified and captured a majority of abnormal activities within the network, further validating the positive impact of SVM-SS on the effectiveness and efficiency of the IDS.

In future work, we plan to enhance the efficiency of the SS algorithm using various selection schemes, integrate other recent optimization algorithms, and use different recent datasets such as DARPA 98/99 and KDD 99.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Feras E. AbuAladas: Conceptualization, Software, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Mohammad Shehab:** Conceptualization, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Rasha Israwah:** Methodology, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Ghaith Jaradat:** Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Yousef Qawqzeh:** Software, Investigation, Writing – original draft, Writing – review & editing.

References

[1] Alsalibi, A. I., Shambour, M. K. Y., Abu-Hashem, M. A., Shehab, M., & Shambour, Q. (2021). Internet of Things in health care: A survey. In A. K. Bhoi, P. K. Mallick, M. N. Mohanty & V. H. C. Albuquerque (Eds.), *Hybrid artificial intelligence and IoT in healthcare* (pp. 165–200). Springer. https://doi.org/10.1007/978-981-16-2972-3_9

[2] Ahmad, W., Rasool, A., Javed, A. R., Baker, T., & Jalil, Z. (2022). Cyber security in IoT-based cloud computing: A comprehensive survey. *Electronics*, *11*(1), 16. https://doi.org/10.3390/electronics11010016

[3] Ngueajio, M. K., Washington, G., Rawat, D. B., & Ngueabou, Y. (2023). Intrusion detection systems using support vector machines on the KDDCUP'99 and NSL-KDD datasets: A comprehensive survey. In *Intelligent Systems and Applications: Proceedings of the 2022 Intelligent Systems Conference, 2*, 609–629. https://doi.org/10.1007/978-3-031-16078-3_42

[4] Alfatemi, A., Rahouti, M., Hsu, D. F., Schweikert, C., Ghani, N., Solyman, A., & Assaqty, M. I. S. (2025). Identifying distributed denial of service attacks through multi-model deep learning fusion and combinatorial analysis. *Journal of Network and Systems Management*, *33*(1), 8. https://doi.org/10.1007/s10922-024-09882-0

[5] Hasan, S. S., & Eesa, A. S. (2020). Optimization algorithms for intrusion detection system: A review. *International Journal of Research-GRANTHAALAYAH*, *8*(08), 217–225. https://doi.org/10.29121/granthaalayah.v8.i8.2020.1031

[6] Glover, F., Laguna, M., & Martí, R. (2003). Scatter search. In A. Ghosh & S. Tsutsui (Eds.), *Advances in evolutionary computing: Theory and applications* (pp. 519–537). Springer. https://doi.org/10.1007/978-3-642-18965-4_20

[7] Antony Vigil, M. S., Ganesh, S., Reddy, P. C., & Babu, R. G. (2024). Interpretable and proactive intrusion detection using discrete optimization learning: Futuristic approach. *Educational Administration: Theory and Practice*, *30*(4), 6668–6681. https://doi.org/10.53555/kuey.v30i4.2461

[8] Kalra, M., Tyagi, S., Kumar, V., Kaur, M., Mashwani, W. K., Shah, H., & Shah, K. (2021). A comprehensive review on scatter search: Techniques, applications, and challenges. *Mathematical Problems in Engineering*, *2021*(1), 5588486. https://doi.org/10.1155/2021/5588486

[9] AlShorman, A., Shannaq, F., & Sheha, M. (2024). Machine learning approaches for enhancing smart contracts security:

A systematic literature review. *International Journal of Data and Network Science*, *8*(3), 1349–1368. https://doi.org/10.5267/j.ijdns.2024.4.007

[10] Abualigah, L., Elaziz, M. A., Shehab, M., Alomari, O. A., Alshinwan, M., Alabool, H., & Al-Arabiat, D. A. (2021). Hybrid Harris Hawks optimization with differential evolution for data clustering. In D. Oliva, E. H. Houssein, & S. Hinojosa (Eds.), *Metaheuristics in machine learning: Theory and applications* (pp. 267–299). Springer. https://doi.org/10.1007/978-3-030-70542-8_12

[11] Rajasoundaran, S., Prabu, A., Kumar, G. S., Malla, P. P., & Routray, S. (2021). Secure opportunistic watchdog production in wireless sensor networks: A review. *Wireless Personal Communications*, *120*(2), 1895–1919. https://doi.org/10.1007/s11277-021-08542-9

[12] Nathanael, K., Cheng, S., Kovalchuk, N. M., Arcucci, R., & Simmons, M. J. (2023). Optimization of microfluidic synthesis of silver nanoparticles: A generic approach using machine learning. *Chemical Engineering Research and Design*, *193*, 65–74. https://doi.org/10.1016/j.cherd.2023.03.007

[13] Hossein, F., Errigo, M., Cheng, S., Materazzi, M., Lettieri, P., Arcucci, R., & Angeli, P. (2025). Acoustic emission and machine learning algorithms for particle size analysis in gas-solid fluidized bed reactors. *Particuology*, *101*, 155–165. https://doi.org/10.1016/j.partic.2024.10.005

[14] Cheng, S., Jin, Y., Harrison, S. P., Quilodrán-Casas, C., Prentice, I. C., Guo, Y. K., & Arcucci, R. (2022). Parameter flexible wildfire prediction using machine learning techniques: Forward and inverse modelling. *Remote Sensing*, *14*(13), 3228. https://doi.org/10.3390/rs14133228

[15] Türk, F. (2023). Analysis of intrusion detection systems in UNSW-NB15 and NSL-KDD datasets with machine learning algorithms. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, *12*(2), 465–477. https://doi.org/10.17798/bitlisfen.1240469

[16] Gaye, B., Zhang, D., & Wulamu, A. (2021). Improvement of support vector machine algorithm in big data background. *Mathematical Problems in Engineering*, *2021*(1), 5594899. https://doi.org/10.1155/2021/5594899

[17] Pan, J. S., Fan, F., Chu, S. C., Zhao, H. Q., & Liu, G. Y. (2021). A lightweight intelligent intrusion detection model for wireless sensor networks. *Security and Communication Networks*, *2021*(1), 5540895. https://doi.org/10.1155/2021/5540895

[18] Akhtar, M. S., & Feng, T. (2021). Deep learning-based framework for the detection of cyberattack using feature engineering. *Security and Communication Networks*, *2021*(1), 6129210. https://doi.org/10.1155/2021/6129210

[19] Belgrana, F. Z., Benamrane, N., Hamaida, M. A., Chaabani, A. M., & Taleb-Ahmed, A. (2021). Network intrusion detection system using neural network and condensed nearest neighbors with selection of NSL-KDD influencing features. In *2020 IEEE International Conference on Internet of Things and Intelligence System*, 23–29. https://doi.org/10.1109/IoTaIS50849.2021.9359689

[20] Nimbalkar, P., & Kshirsagar, D. (2021). Feature selection for intrusion detection system in Internet-of-Things (IoT). *ICT Express*, *7*(2), 177–181. https://doi.org/10.1016/j.icte.2021.04.012

[21] Amaran, S., & Mohan, R. M. (2021). Intrusion detection system using optimal support vector machine for wireless sensor networks. In *2021 International Conference on Artificial Intelligence and Smart Systems*, 1100–1104. https://doi.org/10.1109/ICAIS50930.2021.9395919

[22] Nugroho, E. P., Djatna, T., Sitanggang, I. S., Buono, A., & Hermadi, I. (2020). A review of intrusion detection system in IoT with machine learning approach: Current and future research. In *6th International Conference on Science in Information Technology*, 138–143. https://doi.org/10.1109/ICSITech49800.2020.9392075

[23] Nancy, P., Muthurajkumar, S., Ganapathy, S., Santhosh Kumar, S., Selvi, M., & Arputharaj, K. (2020). Intrusion detection using dynamic feature selection and fuzzy temporal decision tree classification for wireless sensor networks. *IET Communications*, *14*(5), 888–895. https://doi.org/10.1049/iet-com.2019.0172

[24] Kim, J., Kim, J., Kim, H., Shim, M., & Choi, E. (2020). CNN-based network intrusion detection against denial-of-service attacks. *Electronics*, *9*(6), 916. https://doi.org/10.3390/electronics9060916

[25] Ghorbani, A. A., Lu, W., & Tavallaee, M. (2010). *Network intrusion detection and prevention: Concepts and techniques*. USA: Springer. https://dx.doi.org/10.1007/978-0-387-88771-5

[26] Meraihi, Y., Gabis, A. B., Mirjalili, S., & Ramdane-Cherif, A. (2021). Grasshopper optimization algorithm: Theory, variants, and applications. *IEEE Access*, *9*, 50001–50024. https://doi.org/10.1109/access.2021.3067597

[27] Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H. (2021). The arithmetic optimization algorithm. *Computer Methods in Applied Mechanics and Engineering*, *376*, 113609. https://doi.org/10.1016/j.cma.2020.113609

[28] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, *69*, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007

[29] Meng, X., Jiang, J., & Wang, H. (2021). AGWO: Advanced GWO in multi-layer perception optimization. *Expert Systems with Applications*, *173*, 114676. https://doi.org/10.1016/j.eswa.2021.114676

[30] Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: An overview. *Soft Computing*, *22*(2), 387–408. https://doi.org/10.1007/s00500-016-2474-6

How to Cite: AbuAladas, F. E., Shehab, M., Israwah, R., Jaradat, G., & Qawqzeh, Y. (2026). A New Hybrid Approach for Improving Intrusion Detection System Based on Scatter Search Algorithm and Support Vector Machine. *Journal of Computational and Cognitive Engineering*, *5*(1), 143–151. https://doi.org/10.47852/bonviewJCCE52024492