ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
http://dx.doi.org/10.70102/afts.2025.1834.1182

MEDIAN ATTRIBUTE HYBRID CLUSTERED MODEL USING PARTICLE SWARM OPTIMIZATION FOR NETWORK INTRUSION DETECTION

Rajasekhar Kaseebhotla[1*], Dr.K. Raghava Rao[2], Dr. Mallikarjuna Rao[3]

[1]Research Scholar, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India.*
e-mail: rs.kaseebhotla@gmail.com, orcid: https://orcid.org/0000-0001-9004-4469
[2]Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India.
e-mail: raghavarao@kluniversity.in, orcid: https://orcid.org/0000-0003-2119-1275
[3]Professor, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Telangana, India. e-mail: cmrao@giet.ac.in, orcid: https://orcid.org/0000-0003-0674-4353

Received: October 09, 2025; Revised: November 14, 2025; Accepted: December 15, 2025; Published: December 30, 2025

SUMMARY

The rapid growth of cloud-based and large-scale network infrastructures has increased the complexity and frequency of cyber-attacks, demanding efficient and scalable intrusion detection systems (IDS). This paper will also attempt to create a better Network Intrusion Detection System (NIDS) by incorporating a Hybrid Median Attribute Clustering model with Particle Swarm Optimization (MAHC-PSO) to achieve better detection accuracy, less false alarms, and better computation efficiency in high dimensional network space. The suggested MAHC-PSO model utilizes Information Gain in the feature selection on the KDD Cup'99 dataset to minimize dimensions without losing important network features. The median attribute analysis is used to make data representation in a way that is effective and hybrid hierarchical clustering is used to cluster network traffic patterns. Clustering quality is optimized with the help of particle Swarm Optimization that improves the position of the particles depending on the accuracy of detection and the degree of compactness of the cluster. The performance assessment is done with different network sizes between 10, 000 and 60,000 nodes and compared with the current NWFSF-IDMLM and HO-CNN-LSTM-IDS models. The experimental findings indicate that MAHC-PSO model is always better in all the metrics than the benchmark models. The accuracy of feature extraction, hybrid clustering, and cluster set generation were 98.5%, 98.2% and 98.6%, respectively, with 60,000 nodes. PSO fitness value estimation rate was at its highest of 98.8 and median attribute estimation time was lower, at 18 units and pattern analysis time was lower, at 12 units. The overall intrusion detection accuracy was 98.6 which is very high compared to models. The MAHC-PSO model provides a scalable, efficient, and powerful intrusion detection system in large networks in real-time. Its superior accuracy and reduced processing time make it suitable for deployment in cloud-based, enterprise, and future distributed IT environments.

Key words: *clustering, network data, particle swarm optimization, intrusion detection, similar values, median attributes.*

INTRODUCTION

Protecting cloud infrastructure from intruders has become a critical function of intrusion detection systems (IDS), which are essential components of network security. Network Intrusion Detection Systems (NIDS) are used to check on the threats on a number of networks whereas Host Intrusion Detection Systems (HIDS) are used to check on the intrusion on the host. Older systems of IDS were based on misuse detectors, a signature-based model, whereas newer systems are based on anomaly detectors which are built on machine learning (ML) [1]. NIDS are also charged with the responsibility of operating on large amounts of data with a high-dimensional feature set and hence feature selection is an essential step in increasing the accuracy and efficiency of machine learning algorithms. This however poses major problems in processing the huge amount of network data which must be processed in real-time [2],[3].

There are two major types of intrusion detection methods, which are anomaly detection and abuse detection. Anomaly works as a model to detect regular data flow and any alterations are considered as possible intrusion and peculiar in particular to identify an attack that has never been observed before [4]. On the contrary, the abused detection relies on a previously defined model of known intrusions and recognises the unwanted behaviours in the network traffic. In comparison with the old security systems like firewalls, NIDS have an additional advantage of being able to monitor in real-time, gathering data packets, retrieving their properties, and comparing them to known attack signatures. Nevertheless, NIDS have a number of major constraints, including very high rate of false-positive, consumption of resources, failure to detect unknown threats, and dependence on human factors [5].

The goal of the research is to recommend a better way of intrusion detection by introducing a better Median Attribute Hybrid Clustered Model with Particle Swarm Optimization (PSO) to enhance accuracy and efficiency of NIDS [6]. The most important issue is to optimize feature selection and clustering algorithms in order to overcome the problem of false positives and consumption of computer resources as usually encountered in the traditional intrusion detection methods. The suggested model is aimed to consider the scalability and flexibility of intrusion detection system and make them more efficient and reliable in practice [7][8]. This study uses PSO, which is based on the group dynamics of animal behavior, in order to enhance the clustering of network data as an indication of intrusion to improve the overall performance of NIDS.

Although the field of machine learning and techniques of data mining have made great progress towards intrusion detection, the existing systems are still faced with challenges of high false-positive rates and inefficient computation [9]. This is a significant shortcoming of both anomaly-based and abuse-based IDS because detecting hitherto unknown threats is not possible. The conventional IDS systems are highly dependent on set security rules where the system can no longer respond to novel and emerging threats. Furthermore, the performance of existing models is also hampered by the addition of the computational complexity of the processing of large quantities of network data and the inability to select a useful feature. Whereas methods, such as Particle Swarm Optimization (PSO), have demonstrated potential in optimising feature selection, the use of such methods on the intrusion detection systems has not been exhaustively exhausted especially in the presence of hybrid clustering models with potential in handling high dimensions of network data [10][11].

The study hypothesis is that combining Particle Swarm Optimization with a Median Attribute Hybrid Clustered Model would greatly enhance the intrusion detection system to a certain extent by increasing the performance of the systems in terms of feature selection and the clustering capability [12][13]. In particular, the proposed model is expected to minimize false positives, maximize the use of resources, and enhance detection accuracy, especially in the process of occurring unknown threats. The model will improve the general efficiency and scalability of NIDS by capitalizing on the collaborative aspect of PSO particles which in parallel mines patterns on both individual and collective knowledge [14]. The intention of using PSO in this context is to overcome the drawbacks of the traditional methods to provide a more versatile and effective solution to detection of intruders into the IT networks, especially real-time [15].

This study makes the following key contributions:

- Innovative Clustering Model: This is a combination of median attribute selecting and hybrid clustering to obtain precise network data classification.
- PSO Optimization: This employs Particle Swarm Optimization (PSO) to optimize feature selection and clustering resulting in enhanced efficiency and accuracy.
- Minimization of False Positives: Result In minimization of false positives through data clustering and optimization of features.
- Resource Efficiency: Reduces computation costs, ensuring that the model is resource-efficient in large scale intrusion detection.
- Performs better with intrusion detection and is less prone to the false alarm rate than currently available machine learning (ML) and deep learning (DL) models.
- Improved Scalability: Built to support large amounts of data, which provides applicability to real-world IT infrastructures.
- Swarm Intelligence application: This application applies the principles of swarm intelligence using PSO to enhance performance of intrusion detection.

There are five major sections in this article. Section 1 presents the necessity of network intrusion detection, provides an overview of current limitations of the IDS system and introduces the motivation and the contribution of the suggested MAHC-PSO model. Section 2 examines relative literature on machine learning, swarm intelligence and clustering-based intrusion detection methods. Section 3 outlines the suggested methodology, which consists of an attribute selection process, a hybrid of medians and attributes clustering, and PSO optimality, as well as steps of the algorithm. Section 4 provides the results of the experiment and the analysis of performance through several evaluation metrics. Lastly, Section 5 summarizes the study and gives discussions on the future research direction.

LITERATURE SURVEY

Direct power system measurement data is difficult to use in intrusion detection because of the high-dimensionality and large noise levels. Conventional machine learning (ML) methods have a tendency of considering feature processing as preprocessing and therefore they provide erroneous features during training. To solve this, a binary particle swarm-wrapped feature selection methodology (BPSWO) with an improved transfer function, chaotic transformation, and hamming distance was proposed by Han et al. [1] as a solution to the problem of premature convergence in particle swarm optimization. Intrusion detection is then done with the trained classifier after the selection of the features. Network security has become an important component of the contemporary computing due to the increased prevalence of cyber-attacks. Intrusion Detection Systems (IDS), especially Network Intrusion Detection Systems (NIDS), are essential for cybersecurity. However, current systems suffer from limitations like reduced usability, poor endurance, and declining detection accuracy. Deore et al. [2] evaluated the different security measures and suggested the enhancing of such challenges.

Due to the increasing network attacks, there is an increasing need to have effective NIDS to detect malicious activities. The development of machine learning techniques has proved to be significant in the NIDS development because it can process high amounts of data and differentiate between normal and abnormal behavior. Nevertheless, without the appropriate discrimination, deriving features out of large sets of data can raise the level of complexity. To address this, Alsaleh et al. [3] proposed a feature selection technique that extracts the most relevant features, improving detection efficiency and reducing computational overhead. One optimization technique, the Slap Swarm Algorithm (SSA), effectively resolves optimization challenges during feature selection.

Jianhua et al. [4] proposed Flamingo Search Method (FSA) that was inspired by the nomadism behavior of the flamingo, which maximizes global searching and local exploitation. FSA demonstrated strong performance with regards to optimizing complicated problems in all test conditions. Tang et al. [5] introduced IPSO-IRELM, an improved version of the Extreme Learning Machine (ELM), which addresses random initialization issues and adapts dynamically to production networks. This approach enhances the detection system's adaptability.

Injadat et al. [6] suggested NIDS multi-stage ML architecture to balance the performance detection and the complexity of computing. Their analysis involved the analysis of different oversampling methods and feature selection methods to enhance accuracy. Donkol et al. [7] implemented ELSTM together with RNN and LPPSO to overcome the problem of gradient vanishing, overfitting and generalization, as presented in the NSL-KDD data. In a hybrid IDS, Wu et al. [8] used WSVM in conjunction with Deep Belief Networks (DBN), which exhibited better accuracy compared to other traditional ML methods. Salinas et al. [9] utilized multi-objective optimization methods with swarm intelligence to enhance NIDS. Taher et al. [10] introduced a new IoT-based intrusion detection system with the help of tunicate swarm and LSTM. The developments of the ML and swarm optimization emphasize the significance of effective feature selection and computational optimization in the real-world. The insights will be used to advocate the suggestion of Median Attribute Hybrid Clustered Model, which improves the accuracy, scalability, and efficiency of intrusion detection systems.

PROPOSED METHOD

Overview of Clustering in Intrusion Detection

The philosophy of clustering has received much popularity due to its possible aid in intrusion detection systems (IDS). The main objective of clustering is to have the similar data points in a cluster so that the data items in the same cluster are more similar than those in other clusters. This is an advantageous unsupervised learning framework especially in instances where the labels of the data items in classes are not priori. Clustering, through unsupervised learning, helps identify previously unseen attacks on networks. It enables network security experts to label network traffic as benign or malicious without needing expert labeled data, thereby simplifying the laborious task of labeling large scale data sets [15].

Feature Selection and Data Preprocessing

The intrusion detection system is optimized by using the feature selection in order to decrease the dimensions of the data set without losing the necessary information. The KDD Cup'99 is used, which includes 41 features in it, including; src_byte, dsc_byte, duration, and protocol type. Information Gain technique is used in the process of feature selection where features that are redundant and irrelevant are removed [16]. Once the relevant features are chosen, the background noise is removed to prepare the data to be clustered. The purpose of such a preprocessing step is to guarantee that only significant attributes are utilized in the clustering algorithm [17].

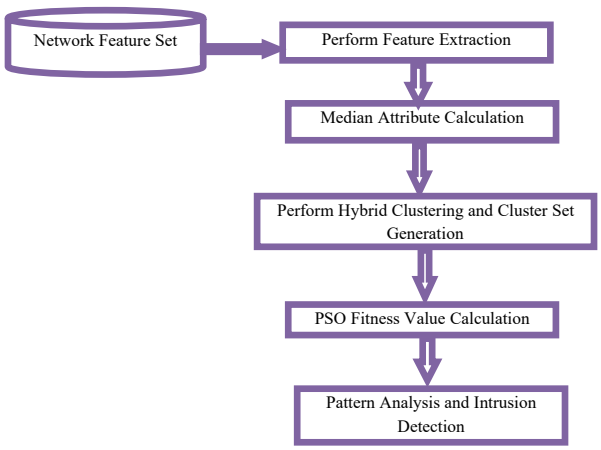

Figure 1. Proposed model framework

The Figure 1 shows the Median attribute Hybrid Clustering (MAHC) methodology combined with Particle Swarm Optimization (PSO) that will be applied in efficient intrusion detection. The framework shows how the network feature set is processed through feature extraction, clustering, and optimization steps, resulting in the generation of the Intrusion Detection Set (IDSset).

Clustering Methodology

The clustering algorithm begins by dividing the dataset into smaller clusters. Each cluster is then analyzed using a predetermined similarity metric to form a hierarchical structure. This hierarchical clustering is required when the data is mixed dataset as the direct mathematical operations such as addition or averageness may not apply [18][19]. The objective of the clustering algorithm is to detect patterns in the data which can be an indication of intrusion attempts. The clustering technique applied in this study is the Median Attribute Hybrid Clustering technique that guarantees the similarity of the data in the study [20]. The clustering process is guided by the following mathematical formulations:

Cluster Size Calculation:

The average size of each cluster is calculated as:

$$Avg\ Size = \frac{\sum_{i=1}^{K} C_i}{K} \qquad (1)$$

In equation (1) C_i represents the number of data points in cluster i, and K is the number of clusters.

Cluster Range Calculation:

Each cluster is characterized by a range, calculated as:

$$\beta = \frac{max(C_i) - min(C_i)}{mean\ range\ of\ cluster} \qquad (2)$$

In equation (2) $max(C_i)$ and $min(C_i)$ represent the maximum and minimum values in the cluster, respectively.

Cluster Grouping Calculation:

Clusters are grouped based on the distance between values, using a distance metric λ. The grouping is performed as:

$$Cluster\ Grouping = maxrange()\ and\ minrange() \qquad (3)$$

In the equation (3) the values greater than the average are grouped into one set, and those lesser than the average are grouped into another.

Particle Swarm Optimization (PSO) for Clustering Optimization

The clustering process is optimized with the help of the Particle Swarm Optimization (PSO). Particles in this optimization method are candidate solutions, whereby each solution has a position and velocity [21][22]. PSO works around these particles to find the best solution. The PSO objective of this study is to improve the clustering process by modifying the particle positions to attain an enhanced performance of the clustering [23].

The process is outlined as follows:

Initialization:

Each particle starts with a random position and velocity in the search space. The position represents a potential solution to the clustering problem.

Fitness Evaluation:

The fitness function evaluates the quality of each particle's solution. The fitness function is based on the accuracy of intrusion detection and the compactness of the clusters.

Particle Update:

The position and velocity of each particle are updated using the following formulas:

$$V_j = V_j + \phi_1 \cdot \left(P_{best} - P_j\right) + \phi_2 \cdot \left(G_{best} - P_j\right) \quad (4)$$

$$P_j = P_j + V_j$$

In equation (4), V_j is the velocity, P_j is the position, P_{best} is the local best position, and G_{best} is the global best position.

Convergence:

The particle swarm iterates until a termination condition (such as the maximum number of iterations, MIt) is met, with the goal of finding the best solution (global optimum) for the clustering problem.

Intrusion Detection and Pattern Analysis

After the process of clustering is done and optimal clusters have been determined, we develop the Intrusion Detection Set (IDSset). The IDSset is a set of network traffic, which is identified as benign or malicious based on the results of the clustering. The data pattern analysis is employed to track and identify inconsistencies in the network traffic data, which is helpful to identify the possible security breaches [24]. The data pattern analysis involves the detection of abnormality in the clustered data patterns, which might be signs of intrusion. This measure will enable real-time identification and categorization of network traffic and further promote the entire system safety [25].

Time Complexity and Scalability

The MAHC-PSO model is designed to be computationally efficient and scalable for handling high-dimensional network traffic data. Feature selection reduces the data dimensionality, which improves processing speed. The hybrid clustering approach ensures that the clustering process is both accurate and efficient. The PSO-based optimization mechanism further enhances the model's scalability, making it suitable for deployment in real-time IT environments. The ability to identify previously unseen attacks without requiring labeled data makes this model robust against evolving intrusion patterns and noise.

Algorithm 1: Median Attribute Hybrid Clustering with Particle Swarm Optimization (MAHC-PSO)

Input: Network Feature Set $\{NF_{set}\}$

Output: Intrusion Detection Set $\{IDS_{set}\}$

Step 1: Feature Extraction

Perform dimensionality reduction by selecting key features from the network data.

Use getattr() to retrieve attributes and simm() model to calculate similarities.

Find high-range attributes using maxrange().

Step 2: Mean Attribute Calculation

Calculate the mean of each attribute to establish threshold values for intrusion detection.

Step 3: Hybrid Clustering

Divide the dataset into smaller clusters using a predetermined similarity metric.

Group clusters based on distance range using maxrange() and minrange().

Step 4: Particle Swarm Optimization (PSO)

Initialize particle positions and velocities.

While iterations < max iterations (MIt):

For each particle Pj:

Calculate fitness function based on cluster accuracy and compactness.

If fitness value > Pbest:

Update Pbest with the current fitness value.

Update particle velocity and position.

Return the best fitness value (gbest).

Step 5: Intrusion Detection and Pattern Analysis

Analyze the generated clusters and detect anomalies.

Generate the intrusion detection set (IDSset) based on clustered data patterns.

The MAHC-PSO algorithm 1 integrates feature extraction, hybrid clustering and Particle Swarm Optimization (PSO) to facilitate effective intrusion detection. It dimensionality data of the network, clusters data of the network traffic, optimizes clustering with the help of PSO, and identifies the anomalies. It is computationally efficient, scalable, and robust, and can be used in detecting intrusions in networks in real-time.

RESULTS

Network Security Overview

The issue of network security has also become prominent, and the primary concern has become the means and the tools used to secure the safety of the networks. Intrusion detection systems (IDS) are at the forefront of this effort, designed to identify illicit usage of networks or computer equipment. An intrusion refers to any activity that compromises the confidentiality, availability, or integrity of a system. As a matter of fact, IDS tracks the various security policy infractions, including unauthorized file or network access. Having a professional understanding of known patterns of intrusions, IDS characterizes user activity as normal or abnormal and gives alerts. However, many traditional IDS approaches suffer from poor detection rates and high false alarm rates. This research paper suggests a hybrid form of data mining model that integrates feature selection, filtering, clustering, divide-and-merge, and ensemble clustering to achieve precision in the detection of intrusions with a false alarm. Median Attribute Hybrid Clustered Model through Particle Swarm Optimization (MAHC-PSO) is a better model compared to the

existing IDS model since it proves to be more effective in detecting network intrusions. This performance advantage is essential in real-time IT intrusion detection, which underscores the appropriateness of the model for deployment in large-scale networks, herein the cloud-based Network Intrusion Detection Systems (NIDS).

Feature Extraction Accuracy

The process of feature extraction is very important in intrusion detection. It entails the manipulation of the unprocessed data into useful features, which can be utilized by machine learning algorithms to perform classification. In feature extraction, a more cost and time-efficient smaller set of features (F) is chosen among a larger set (N). By removing non-dominant features, we reduce the training time and improve model simplicity. Table 1 shows the levels of accuracy of the feature extraction of the current model and a proposed model, which reveals that the MAHC-PSO model has the highest level of accuracy in comparison with other existing methods.

Table 1. Feature extraction accuracy levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	97.4	93.0	94.1
20000	97.6	93.2	94.4
30000	97.9	93.5	94.7
40000	98.1	93.7	94.9
50000	98.3	93.9	95.0
60000	98.5	94	95.2

The table presents the feature extraction accuracy of different intrusion detection models under varying network sizes. As the number of nodes increases from 10,000 to 60,000, the proposed MAHC-PSO model consistently achieves higher accuracy compared to the NWFSF-IDMLM and HO-CNN-LSTM-IDS models. This demonstrates the robustness and scalability of the MAHC-PSO approach in effectively extracting relevant features from large-scale network traffic, which contributes to improved intrusion detection performance.

Median Attribute Calculation Time

This table shows the feature extraction accuracy of the various intrusion detection models with varying network sizes. The proposed MAHC-PSO model has a higher accuracy than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models as the number of nodes is increased between 10,000 and 60,000. It shows how the MAHC-PSO approach can be used to extract the relevant features in large-scale network traffic with high effectiveness and scalability, which is a factor in enhancing intrusion detection performance.

Table 2. Median attribute calculation time levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	17.0	23.0	27.0
20000	17.2	23.2	27.2
30000	17.3	23.4	27.4
40000	17.6	23.6	27.6
50000	17.8	23.8	27.9
60000	18	24	28

Table 2 displays the median attribute calculation time of various intrusion detection models of the network of increasing size. In comparison to NWFSF-IDMLM and HO-CNN-LSTM-IDS models, the proposed MAHC-PSO model is always much faster to compute based on the number of nodes in the network. This saves in processing time, indicating that the MAHC-PSO method is more efficient in

computation time, and it is thus more appropriate for large-scale and real-time network intrusion detection functions.

Hybrid Clustering Accuracy

Clustering is a method that is used to group data in terms of similarity. The hybrid clustering proposal that is implemented in intrusion detection separates the dataset into small clusters and constructs hierarchies to help in the proper detection of abnormal behavior. The Table 3 results indicate that the MAHC-PSO model performs better than the already existing models in the accuracy of clustering and shows that this model can be used to facilitate better detection of complex intrusion patterns in enterprise IT networks.

Table 3. Hybrid clustering accuracy levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	97.0	94.1	92.4
20000	97.3	94.3	92.7
30000	97.5	94.5	92.9
40000	97.8	94.7	93.1
50000	98.0	94.8	93.3
60000	98.2	95	93.5

Cluster Set Generation Accuracy

Another important factor of the clustering process is cluster set generation, the aim of which is to find and isolate odd data points. Table 4 demonstrates that the MAHC-PSO model is more effective in the formation of accurate clusters. This is essential in improving the accuracy and speed of the intrusion detection systems.

Table 4. Cluster set generation accuracy levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	97.7	94.0	94.8
20000	97.9	94.2	94.9
30000	98.0	94.4	95.1
40000	98.2	94.5	95.4
50000	98.4	94.8	95.6
60000	98.6	95	95.8

Table 4 is used to compare the accuracy of the cluster set generation of various intrusion detection models at varying network sizes. As the number of nodes grows, the suggested MAHC-PSO model is more accurate than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models. It means that the MAHC-PSO technique is better suited to construct precise and valid cluster sets, which strengthens the general capacity to identify the intrusion in large-scale environments.

PSO Fitness Value Calculation Accuracy

The quality of potential solutions is measured with the help of the Particle Swarm Optimization (PSO) fitness value. Table 5 depicts that the MAHC-PSO model has an impressive enhancement in calculating the fitness values. The optimization is essential in the efficient detection of the most pertinent intrusion detection features.

Table 5. PSO fitness value calculation accuracy levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	98.0	94.0	93.5
20000	98.1	94.2	93.6
30000	98.3	94.5	93.9
40000	98.5	94.6	94.1
50000	98.7	94.9	94.3
60000	98.8	95	94.6

The table shows that PSO fitness value computation is accurate with varied intrusion detection models and varying network sizes. With an increasing number of nodes, the proposed MAHC-PSO model regularly gains greater fitness accuracy than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models. This enhanced fitness assessment is a measure of how the PSO-based optimization has succeeded in optimizing cluster quality and overall performance of the intrusion detection system, especially in large-scale network settings.

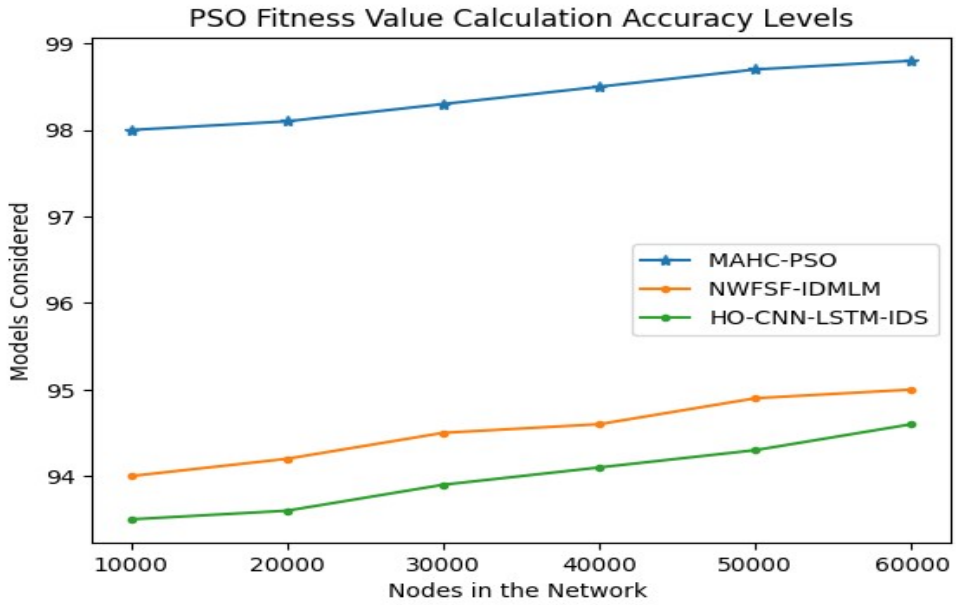

Figure 2: PSO fitness value calculation accuracy levels

Figure 2 shows the accuracy of the calculation of the PSO fitness value of the various intrusion detection models with the growth in the size of the network. The presented MAHC-PSO model is always more accurate in its fitness than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models. The gradual increasing nature implies that the PSO-based optimization is effective and scalable to enhance the performance of clustering in detecting the network intrusion of a large-scale network.

Pattern Analysis Time

Clustering is a method that clusters the data based on similarities. The suggested hybrid clustering system, used in the intrusion detection area, separates the data set into smaller groups and constructs the hierarchical schemes for proper abnormal behavior recognition. As indicated in Table 6, the MAHC-PSO model performs better than the existing models in clustering accuracy, and it can be said that better performance can be achieved by the model in detecting complex patterns of intrusion in an enterprise IT network.

Table 6. Pattern analysis time levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	11.1	17.0	22.3
20000	11.3	17.2	22.5
30000	11.5	17.4	22.6
40000	11.6	17.6	22.7
50000	11.8	17.8	22.9
60000	12	18	23

The table shows the pattern analysis time of the various intrusion detection models in relation to the network size. The MAHC-PSO model remains shorter than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models with an increase in the number of nodes of the network. This low processing time indicates the scalability and efficiency of the MAHC-PSO approach, which is very appropriate in real-time intrusion detection of large-scale network environments.

Intrusion Detection Accuracy

Lastly, the accuracy levels of intrusion detection of the proposed MAHC-PSO model are significantly better than those of the current models. Table 7 and Figure 3 indicate that the MAHC-PSO model has the highest accuracy, which is a prerequisite to provide solid security against network intrusions in enterprises.

Table 7. Intrusion detection accuracy levels

Nodes in the Network	Models Considered		
	MAHC-PSO Model	NWFSF-IDMLM Model	HO-CNN-LSTM-IDS Model
10000	97.7	94.3	93.8
20000	97.9	94.5	94.0
30000	98.0	94.7	94.2
40000	98.3	94.9	94.3
50000	98.5	95.0	94.5
60000	98.6	95.2	94.6

It is seen in the table that the intrusion detection ability of the different models improves with the size of the network, and the proposed MAHC-PSO model has better intrusion detection than the other models of NWFSF-IDMLM and HO-CNN-LSTM-IDS at any node count. This enhancement proves that the MAHC-PSO approach is robust and scalable in the detection of intrusion in large network environments.

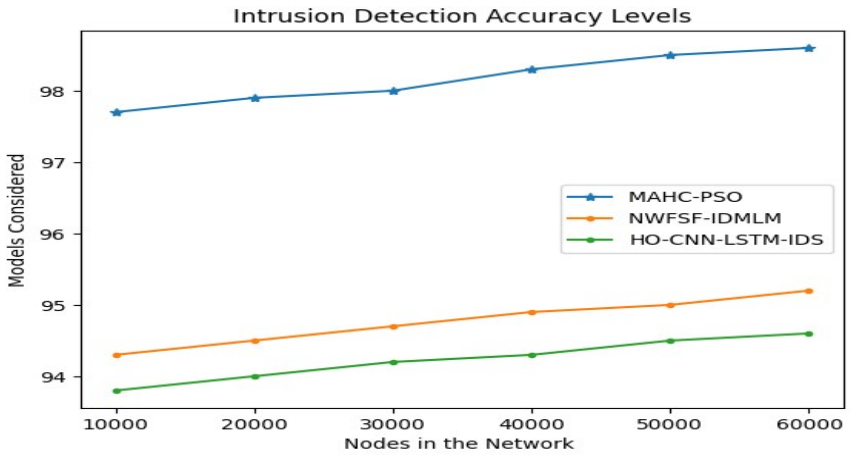

Figure 8. Intrusion detection accuracy levels

The graph illustrates the accuracy of intrusion detection of various models with an enlarging network. The proposed MAHC-PSO model is more effective than the NWFSF-IDMLM and HO-CNN-LSTM-IDS models and is more accurate at all node counts. This growing pattern points to the strength and scalability of the MAHC-PSO methodology for effective intrusion detection in a large-scale network setup.

CONCLUSION

This paper has introduced a Median Attribute Hybrid Clustering model, which is optimized by Particle Swarm Optimization (MAHC-PSO), in order to detect network intrusion effectively. The suggested solution combines feature selection, median-based attribute analysis, hybrid clustering, and PSO-based optimization in order to improve the detection accuracy and minimize the cost of computation. The better performance of the MAHC-PSO model compared to the currently used NWFSF-IDMLM and HO-CNN-LSTM-IDS models was proven by extensive experimental assessment with the use of different network sizes. The findings indicate that the proposed MAHC-PSO model had higher accuracy in feature extraction, with a high of 98.5% of the 60,000 network nodes in the model, and a much lower median attribute calculation time (18 units) than the benchmark models. In addition, the accuracy of the hybrid clustering strategy with the median-based clustering strategy rose to 98.2%, and the generation of cluster sets accuracy rose to 98.6%, which proves the efficiency of the clustering quality in the strategy. Pattern analysis was also shortened by the model down to 12 units, which allowed detection to occur quickly. Above all, it was shown that the MAHC-PSO model has the highest intrusion detection rate, with the highest rate of 98.6%, which is a higher value compared to the current methods of all network sizes, which emphasizes its strength and scalability. In general, the MAHC-PSO model is computationally efficient, accurate, and well-scaled to large-scale and real-time network intrusion detection.

Future research will be dedicated to the attempt to integrate the proposed MAHC-PSO model with cloud-based intrusion detection systems and assess its functionality in real-time distributed IT systems, such as its implementation in enterprise networks, its extension to IoT and distributed IT systems, as well as its hybridization with deep learning models in an attempt to advance the adaptive intrusion detection capabilities further.

Acknowledge

Acknowledgment: The authors declare that they have no conflict of interest. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript. The article has no research involving Human Participants and/or Animals. The author has no financial or proprietary interests in any material discussed in this article.

COMPLIANCE WITH ETHICAL STANDARDS:

Conflicts of Interest:

The authors declare that they have no conflict of interest. The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Availability of data and material:

No data and materials are available for this paper. Data sharing is not applicable to this article, as no datasets were generated or analyzed during the current study.

Ethical Approval:

The article has no research involving Human Participants and/or Animals

Competing Interest:

The author has no financial or proprietary interests in any material discussed in this article.

DECLARATIONS:

Funding:

No Funding is applicable.

Code availability:

The data and code can be given based on the request

REFERENCES

[1] Han Y, Wang Y, Cao Y, Geng Z, Zhu Q. A novel wrapped feature selection framework for developing power system intrusion detection based on machine learning methods. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2023 Aug 2;53(11):7066-76. https://doi.org/10.1109/TSMC.2023.3292110

[2] Deore B, Bhosale S. Hybrid optimization enabled robust CNN-LSTM technique for network intrusion detection. Ieee Access. 2022 Jun 15;10:65611-22. https://doi.org/10.1109/ACCESS.2022.3183213

[3] Alsaleh A, Binsaeedan W. The influence of salp swarm algorithm-based feature selection on network anomaly intrusion detection. IEEe Access. 2021 Aug 3;9:112466-77. https://doi.org/10.1109/ACCESS.2021.3102095

[4] Zhiheng W, Jianhua L. Flamingo search algorithm: a new swarm intelligence optimization algorithm. IEEE Access. 2021 Jun 18;9:88564-82. https://doi.org/10.1109/ACCESS.2021.3090512

[5] Tang Y, Li C. An online network intrusion detection model based on improved regularized extreme learning machine. IEEE Access. 2021 Jun 29;9:94826-44. https://doi.org/10.1109/ACCESS.2021.3093313

[6] Injadat M, Moubayed A, Nassif AB, Shami A. Multi-stage optimized machine learning framework for network intrusion detection. IEEE Transactions on Network and Service Management. 2020 Aug 7;18(2):1803-16. https://doi.org/10.1109/TNSM.2020.3014929

[7] Donkol AA, Hafez AG, Hussein AI, Mabrook MM. Optimization of intrusion detection using likely point PSO and enhanced LSTM-RNN hybrid technique in communication networks. IEEE Access. 2023 Jan 26;11:9469-82. https://doi.org/10.1109/ACCESS.2023.3240109

[8] Wu Y, Lee WW, Xu Z, Ni M. Large-scale and robust intrusion detection model combining improved deep belief network with feature-weighted SVM. IEEE Access. 2020 May 15;8:98600-11. https://doi.org/10.1109/ACCESS.2020.2994947

[9] Salinas O, Soto R, Crawford B, Olivares R. An integral cybersecurity approach using a many-objective optimization strategy. IEEE Access. 2023 Aug 22;11:91913-36. https://doi.org/10.1109/ACCESS.2023.3307492

[10] Taher F, Elhoseny M, Hassan MK, El-Hasnony IM. A novel tunicate swarm algorithm with hybrid deep learning enabled attack detection for secure iot environment. IEEE Access. 2022 Dec 5;10:127192-204. https://doi.org/10.1109/ACCESS.2022.3226879

[11] Ibrahim MS, Dong W, Yang Q. Machine learning driven smart electric power systems: Current trends and new perspectives. Applied Energy. 2020 Aug 15;272:115237. https://doi.org/10.1016/j.apenergy.2020.115237

[12] Sumaiya Thaseen I, Saira Banu J, Lavanya K, Rukunuddin Ghalib M, Abhishek K. An integrated intrusion detection system using correlation-based attribute selection and artificial neural network. Transactions on Emerging Telecommunications Technologies. 2021 Feb;32(2):e4014. https://doi.org/10.1002/ett.4014

[13] Elmasry W, Akbulut A, Zaim AH. Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic. Computer Networks. 2020 Feb 26;168:107042. https://doi.org/10.1016/j.comnet.2019.107042

[14] Wang H, Cao Z, Hong B. A network intrusion detection system based on convolutional neural network. Journal of Intelligent & Fuzzy Systems. 2020 Jun 25;38(6):7623-37. https://doi.org/10.3233/JIFS-179833

[15] Ferrag MA, Maglaras L, Moschoyiannis S, Janicke H. Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. Journal of Information Security and Applications. 2020 Feb 1;50:102419. https://doi.org/10.1016/j.jisa.2019.102419

[16] Al-Jame F, Ali W. Finite Element Evaluation of Mechanical Performance of Hybrid Composite Aerospace Structures. Journal of Reconfigurable Hardware Architectures and Embedded Systems. 2024 Dec 21:12-7.

[17] Jiang H, He Z, Ye G, Zhang H. Network intrusion detection based on PSO-XGBoost model. IEEE Access. 2020 Mar 23;8:58392-401. https://doi.org/10.1109/ACCESS.2020.2982418

[18] Bovenzi G, Aceto G, Ciuonzo D, Persico V, Pescapé A. A hierarchical hybrid intrusion detection approach in IoT scenarios. In GLOBECOM 2020-2020 IEEE global communications conference 2020 Dec 7 (pp. 1-7). IEEE. https://doi.org/10.1109/GLOBECOM42002.2020.9348167

[19] Shyla S, Bhatnagar V, Bali V, Bali S. Optimization of intrusion detection systems determined by ameliorated HNADAM-SGD algorithm. Electronics. 2022 Feb 9;11(4):507. https://doi.org/10.3390/electronics11040507

[20] Powers DM. Evaluation: from Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation. arXiv preprint arXiv:2010.16061. 2020.

[21] Khraisat A, Alazab A. A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. Cybersecurity. 2021 Mar 8;4(1):18. https://doi.org/10.1186/s42400-021-00077-7

[22] Ulkilan A. Optimization of Power Flow in Hybrid Microgrids Using AI-Based Algorithms. Journal of Scalable Data Engineering and Intelligent Computing. 2024 Dec 9;1(1):41-7.

[23] Chaganti R, Varadarajan V, Gorantla VS, Gadekallu TR, Ravi V. Blockchain-based cloud-enabled security monitoring using internet of things in smart agriculture. Future Internet. 2022 Aug 24;14(9):250. https://doi.org/10.3390/fi14090250

[24] Rahman F. Artificial Intelligence-Driven Cybersecurity Framework for Industrial Control Networks. Transactions on Secure Communication Networks and Protocol Engineering. 2025 Mar 5;2(1):1-8.

[25] Li M, Liu Y, Tian Z, Shan C. Privacy protection method based on multidimensional feature fusion under 6G networks. IEEE Transactions on Network Science and Engineering. 2022 Jun 29;10(3):1462-71. https://doi.org/10.1109/TNSE.2022.3186393