International Journal of Engineering Research and Science & Technology

IJERST

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1 (2026)

ijerst.editor@gmail.com
editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 1, 2026

Peer Reviewed, Referred & Indexed Journal

Research Paper

NETWORK ATTACK DETECTION BY GROUP SEARCH OPTIMIZATION USING CONVOLUTIONAL DEEP LEARNING MODEL

Sharjil Iqbal
Senior Data Engineer
Department of Marketing, University Loyola University Chicago
Iqbal.sharjil1987@gmail.com

ABSTRACT: Intrusion Detection Systems (IDS) play a vital role in safeguarding modern networks against increasingly complex cyberattacks. However, the presence of high-dimensional, redundant, and noisy features in network traffic data often degrades detection accuracy and scalability. To address this challenge, this paper proposes GSFOIDS (Group Search Based Feature Optimization for Intrusion Detection System), a hybrid framework that integrates Group Search Optimization (GSO) with a Convolutional Neural Network (CNN) for efficient and accurate intrusion detection. Initially, the intrusion dataset is cleaned and preprocessed to remove irrelevant attributes and improve data quality. GSO is then employed to identify an optimal subset of discriminative features using cooperative producer, scrounger, and ranger search strategies. A CNN-based fitness function guides the optimization process by assessing classification accuracy. Experimental results demonstrate that GSFOIDS significantly outperforms existing models across multiple evaluation metrics

Index Terms- Anomaly Detection, ANN, Clustering, Genetic Algorithm, Intrusion Detection.

1. Introduction

Computer networks represent one of the latest advancements in IT services, offering the advantage of accessibility regardless of location or time. They provide flexibility in adjusting storage capacities, reduce costs, and support mobile and collaborative applications and services. Additionally, network services are multisource, allowing users to choose from different providers based on their specific needs. By utilizing computer networks, organizations can save on storage costs, power consumption, physical space, and maintenance. As these services become increasingly widespread, businesses, banks, and governments are adopting the technology. However, this expansion has also exposed systems to a range of cyberattacks, prompting the need for robust security measures. Many network service providers offer various security features as part of their service packages. For instance, Amazon Web Services (AWS) provides security services with limited validity based on the duration of the service license.

Traditional intrusion detection systems (IDS) are insufficient for managing the vast data flow within modern network infrastructures, which experience high volumes of traffic. Most conventional IDSs are single-threaded, whereas multi-threaded IDSs are required in cloud environments to handle large-scale data flow. In traditional networks, IDSs monitor, detect, and alert administrators to suspicious activities by being deployed at critical points on the user's site. However, in a cloud-based network, IDSs must be installed on the cloud server and operated by the service provider. In cases where an attacker breaches and compromises user data, the cloud user may not be directly informed. The service provider receives the intrusion data and may choose not to disclose the incident to the user, potentially to protect their reputation. To address this, a neutral third-party

monitoring service can offer unbiased oversight, providing alerts and reports to both cloud users and service providers.

This paper proposes a multi-cloud IDS managed by a third-party monitoring service. This service would deliver alert reports and expert advice to both cloud users and providers, ensuring transparency and reliable intrusion detection. The proposed system introduces an efficient and reliable distributed cloud IDS approach to overcome the limitations of traditional IDSs.

II. Related Work

Kabir et al. [7] developed an Optimum Allocation Least Square Support Vector Machine (OALSSVM) model, where "optimum allocation" refers to selecting specific sessions from a dataset to train the SVM model. The accuracy of the OALSSVM model in detecting intrusions depends on these selected sessions, leading to improved performance.

Chuanlong Yin [8] explored an intrusion detection system using a recurrent neural network (RNN-IDS) based on deep learning. The study evaluated the model's performance in both binary and multiclass classification, analyzing the impact of factors like the number of neurons and learning rate. Yin compared the RNN-IDS against other machine learning models such as J48, artificial neural networks, random forests, and SVMs using a benchmark dataset.

Kaiyuan et al. [9] proposed a network intrusion detection algorithm that combines hybrid sampling with a deep hierarchical network. Their method applies one-side selection (OSS) to remove noise from majority samples and SMOTE to increase minority samples, creating a balanced dataset. The model uses a CNN to extract spatial features and BiLSTM for temporal features, enhancing detection accuracy and reducing training time.

In another study [10], the authors compared the performance of neural networks, SVMs, and decision trees for anomaly detection. They found that the effectiveness of machine learning algorithms depends heavily on the practical context. Machine learning approaches like neural networks, SVMs, and decision trees are widely used in anomaly detection to improve classification speed and accuracy. Additionally, techniques such as genetic algorithms and information theory further enhance classification by combining theoretical support.

Zina et al. [11] introduced two models for intrusion detection and classification: the Trust-based Intrusion Detection and Classification System (TIDCS) and its accelerated version, TIDCS-A. TIDCS uses advanced feature selection, grouping features to optimize classification. It maintains system integrity by updating trust relationships between nodes, while TIDCS-A enhances this by dynamically determining optimal times for node maintenance, reducing exposure to attacks. Both models base their final classification decisions on a combination of node behavior history and machine learning algorithms, with detected attacks lowering node trustworthiness.

R. Ben et al. [12] proposed a hybrid approach that combines Convolutional Neural Networks (CNN) with bidirectional Long Short-Term Memory (BiLSTM) networks to enhance network intrusion detection capabilities. This approach supports both binary and multiclass classification tasks effectively. The efficacy of this hybrid model was validated using established datasets like UNSW-NB15 and NSL-KDD, demonstrating superior detection performance.

T. Kim et al. [13] developed a method to discern complex packet patterns for distinguishing between network intrusions and benign sessions. They created a new training dataset for a Generative Adversarial Network (GAN) using misclassified data from an original LSTM-DNN-trained dataset. This GAN assesses whether a received packet can be accurately classified by the LSTM-DNN model. If uncertainty arises in classification, the detection process pauses and retries with the next packet. The refined classification algorithm based on LSTM-DNN and the validation model using GAN enables

precise real-time network intrusion detection without interrupting sessions or requiring delays to accumulate packets.

III. Methodology

The proposed GSFOIDS (Group Search Based Feature Optimization for Intrusion Detection System) model is described in detail in this section. The entire process is divided into two main modules: the training module and the testing module. In the first module, the development of a feature ontology and the training of an CNN network are undertaken. The second module focuses on the testing and evaluation of the trained CNN model. Figure 1 illustrates the operational blocks of the proposed model.

Dataset Cleaning The dataset cleaning stage involves removing unwanted information from the dataset to improve the quality of the data. The input data contains various attributes, each with its specific relevance to the analysis. For instance, the input dataset used in this study consists of several fields, but some initial feature values, such as session ID, connection type, and transferring protocol, were excluded from the analysis [12]. The cleaned dataset is structured into a matrix format with rows and columns, where each row represents a session and each column represents a feature set associated with that session.

$$PD \leftarrow Pre\_Processing(RD)$$

Where ID is intrusion dataset and PID is Preprocessed Intrusion Dataset

Feature Optimization Following preprocessing, the feature matrix is further examined to determine the most informative attributes that directly support intrusion classification. To construct an efficient feature selection framework, this work adopts the Group Search Optimization (GSO) algorithm, which mimics the searching behavior of animals operating collaboratively within a group. The objective of this optimization stage is to reduce redundancy while retaining the most discriminative features for classification.

Group Search Optimization In the proposed approach, an initial population of candidate feature subsets is generated randomly using a Gaussian-based distribution. Each candidate solution represents a search member and is encoded as a binary vector consisting of 0s and 1s, where a value of 1 indicates that a particular feature is selected and a value of 0 signifies its exclusion. This binary representation is similar to chromosome encoding used in evolutionary algorithms, enabling efficient manipulation and evaluation of feature subsets.

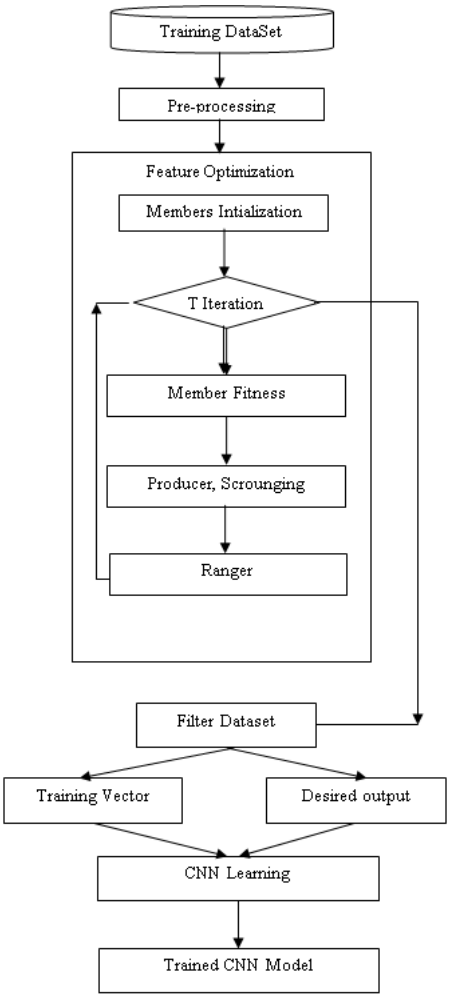

Fig. 1 Block diagram of GSFOIDS.

Within the GSO framework, search members cooperate by exploring the feature space through producer, scrounger, and ranger roles. Producers guide the search toward promising regions, scroungers exploit nearby solutions, and rangers maintain diversity by exploring unexplored areas.

This coordinated search mechanism enables effective exploration and exploitation of the feature space, leading to an optimized feature subset that enhances intrusion detection performance.

Gm Generate_Group_Member(m, n)

Fitness To assess the quality of each candidate solution in the GSO population, a fitness function is defined using a Convolutional Neural Network (CNN). For each feature subset, a temporary CNN model is trained using the selected features, and its classification accuracy is computed. The accuracy reflects the model's ability to correctly identify different intrusion categories present in the dataset. This accuracy value is used as the fitness score for the corresponding group member. A higher fitness score member is Producer, indicates that the selected feature subset contributes more effectively to intrusion classification, thereby guiding the optimization process toward superior solutions.

P←Members_Fitness(Gm)

Group Sprounger Update Strategy Based on fitness evaluation, the best-performing solution within the group is identified as the leader. Other group members update their feature selections by partially adapting the leader's binary pattern, promoting convergence toward optimal solutions. This update mechanism allows promising feature combinations to propagate throughout the group while maintaining sufficient diversity to avoid premature convergence.

Gm←_update(Gm, C)

Ranger After updating the group, solutions with consistently low fitness values are removed from the population. These weak candidates are replaced with newly generated solutions, ensuring continuous exploration of the feature space and preventing stagnation during optimization.

Final Feature After completing a predefined number of iterations, the algorithm selects the optimal feature subset from the final population. In the second phase of the proposed GSFOIDS framework, the raw intrusion dataset is filtered using the optimized feature set obtained through GSO. The selected features are normalized and provided as input to a deep learning model for classification.

F←Filter(PD,P)

Convolution Neural Network The CNN is employed as the core learning model due to its ability to capture structural patterns and reuse shared weight parameters. CNNs utilize convolution and pooling operations to extract hierarchical feature representations while maintaining robustness to geometric transformations in the data. These characteristics make CNNs well suited for intrusion detection tasks involving complex feature interactions [14, 15].

TrainData←TrainingData_Filter(PD,F)

CNN_Model ←Train_CNN(Train_data)

IV. EXPERIMENT AND RESULTS

The proposed GSFOIDS framework, together with the baseline comparison models, was developed and tested using the MATLAB environment. All experimental evaluations were performed on a system configured with an Intel Core i3 processor (6th generation) and 4 GB of RAM. The dataset used for model training and testing was obtained from the source described in [18]. To assess its effectiveness, the performance of the GSFOIDS model was benchmarked against an existing malicious network session detection approach, namely the G-CNN model reported in [9].

Results

Table 1 Precision based comparison of IDS models.

Dataset	Previous Model	GSFOIDS
4000	0.7567	0.9574
8000	0.6154	0.9756
12000	0.6847	0.9683
16000	0.2588	0.9772

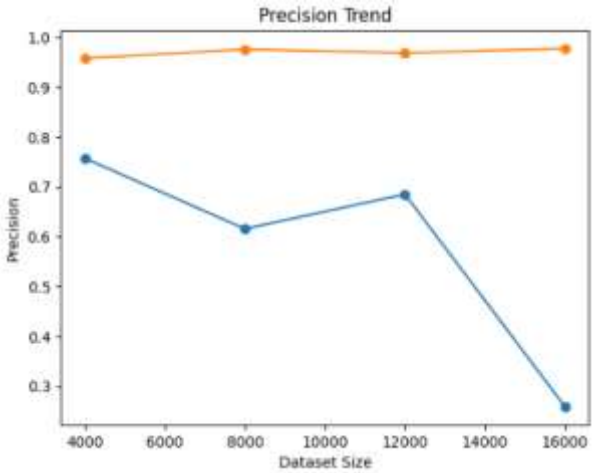

Fig.2 Precision value and testing dataset based comparison of models.

Precision reflects the model's ability to correctly identify attack instances while minimizing false positives. As shown in Table 1, the GSFOIDS model consistently achieves significantly higher precision values than the previous model for all dataset sizes. Notably, while the precision of the previous model degrades sharply as the dataset size increases (dropping to 0.2588 at 16,000 samples), GSFOIDS maintains stable and high precision above 0.95. Fog. 2 demonstrates the robustness of the proposed feature optimization and CNN-based classification approach. The bar and line graphs further emphasize the consistent superiority of GSFOIDS across varying data volumes.

Table 2 Recall based comparison of IDS models.

Dataset	Previous Model	GSFOIDS
4000	1	0.985
8000	1	0.947
12000	0.769	0.9586
16000	0.3278	0.9764

Recall measures the effectiveness of a model in correctly detecting actual attack instances. Although the previous model shows perfect recall for smaller datasets, its performance drops drastically for larger datasets. In contrast, GSFOIDS sustains high recall values across all dataset sizes, indicating reliable detection capability even with increased data complexity. The graphical trends confirm that GSFOIDS provides a more balanced and scalable detection performance.

Table 3 F-measure based comparison of IDS models.

Dataset	Previous Model	GSFOIDS
4000	0.8615	0.971
8000	0.7619	0.9611
12000	0.7244	0.9586
16000	0.2892	0.9768

The F-measure combines both precision and recall, offering a holistic assessment of detection performance. The GSFOIDS model achieves consistently high F-measure values, exceeding 0.95 for all tested datasets. Conversely, the previous model exhibits a steep decline in F-measure as dataset size increases. The bar and line plots clearly illustrate the improved stability and effectiveness of GSFOIDS in handling large-scale network traffic.

Table 4 Accuracy based comparison of IDS models.

Dataset	Previous Model	GSFOIDS
4000	75.67	95.67
8000	61.54	96.99
12000	66.49	97.28
16000	30.96	97.36

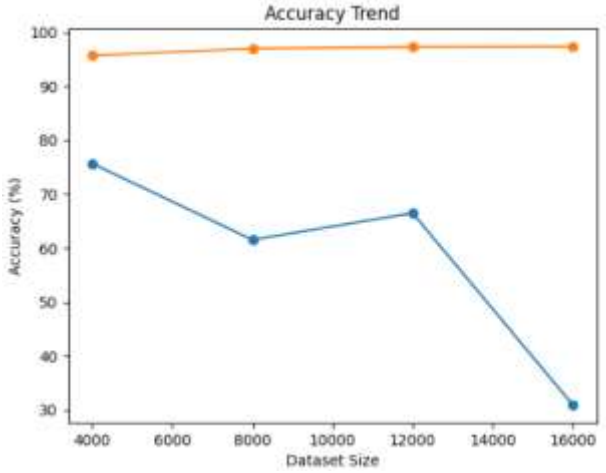

Fig. 3 Accuracy Percentage and testing dataset based comparison of models.

Accuracy represents the overall correctness of attack classification. As shown in Table 4, GSFOIDS achieves accuracy levels above 95% across all dataset sizes, whereas the previous model's accuracy falls below 31% for larger datasets. Fig. 3 substantial improvement highlights the effectiveness of the Group Search Optimization–based feature selection combined with CNN learning. The accuracy graphs further validate the consistent performance gains delivered by the proposed model.

IV. Conclusion

This paper presented GSFOIDS (Group Search Based Feature Optimization for Intrusion Detection System), a hybrid intrusion detection framework that integrates Group Search Optimization (GSO) with a Convolutional Neural Network (CNN) to achieve accurate and scalable network attack detection. The proposed model was structured into two main phases: a training phase focused on feature optimization and CNN learning, and a testing phase dedicated to performance evaluation on unseen data. Experimental results clearly demonstrate the effectiveness of the proposed approach. Across all dataset sizes, GSFOIDS consistently outperformed the previous model in terms of precision, recall, F-measure, and accuracy. While the baseline model showed a significant decline in performance as the dataset size increased—particularly beyond 12,000 sessions—the GSFOIDS model maintained stable and high detection capability. Precision values for GSFOIDS remained above 0.95, recall exceeded 0.94, and F-measure stayed close to 0.97, indicating a strong balance between false positive and false negative detection. Moreover, the classification accuracy of GSFOIDS remained consistently above 95%, even for the largest dataset, highlighting its robustness and scalability.

References

1. Fuhong Lin, Yutong Zhou, Xingsuo An, Ilsun You, Fair Resource Allocation in an Intrusion Detection System: Ensuring the Security of Internet of Things Devices, IEEE conference on Consumer electronics Computing Magazine, Volume: 7, Issue: 6, ISSN: 2162-2248, publishedYear 2018.

2. Mohammad Saeid Mahdavinejad, Mohammadreza Rezvan, Mohammadamin Barekatain Peyman Adibi, Payam Barnaghi, Amit P Sheth Machine learning for internet of things data analysis: a survey Digital Communications and Networks Science Direct, Volume:4, Issue 3, Pages: 161- 175, published Year 2018.

3. Sai Kiran, K.V.V.N.L. R.N. Kamakshi Devisetty, N. Pavan Kalyan, K. Mukundini, R. Karthi. "Building a Intrusion Detection System for IoT Environment using Machine Learning Techniques". Procedia Computer Science, Volume 171, 2020, Pages 2372-2379.

4. Bacem Mbarek, Mouzhi Ge, and Tomás Pitner. 2020. Enhanced nnetwork intrusion detection system protocol for internet of things. In Proceedings of the 35th Annual ACM Symposium on Applied Computing (SAC '20). Association for Computing Machinery, New York, NY, USA, 1156–1163.

5. Ganji, M. (2025). Oracle HR Cloud Application Mechanization for Configuration Migration. International Journal Of Engineering Development And Research, 13(2). https://doi.org/10.56975/ijedr.v13i2.301303

6. Gaddam, S. (2024). Integrating machine learning models with continuous integration and continuous delivery (CI/CD) pipelines for a learning-driven approach to software engineering.
7. Bahram Hajimirzaei and Nima Jafari Navimipour. 2019. Intrusion detection for computer network using neural nnetworks and artificial bee colony optimization algorithm. ICT Express 5, 1 (2019), 56–59.
8. E. Kabir, J. Hu, H. Wang, and G. Zhuo, ‘‘A novel statistical technique for intrusion detection systems,’’ Future Gener. Comput. Syst., vol. 79, pp. 303–318, Feb. 2018.
9. Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines. (2025). Inaternational Journal of Communication Networks and Information Security, 16(4). https://doi.org/10.48047/ijcnis.16.4.2111
10. ChuanlongYin ,Yuefei Zhu, Jinlong Fei, And Xinzheng He. “A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks” current version November 7, 2017.
11. S. M. K. P. (2025). Cryptography in iOS: A Study of Secure Data Storage and Communication Techniques. International Journal on Science and Technology, 16(1). https://doi.org/10.71097/ijsat.v16.i1.1403.
12. Mallick, P. (2020). Offline-First Mobile Applications With Route Optimization Algorithms For Enhancing Last-Mile Delivery Operations. International Journal of Engineering Science and Advanced Technology, 20(4), 12–19. https://doi.org/10.64771/ijesat.2020.v20.i04.pp12-19.
13. Zina Chkirbene, AimanErbad, RidhaHamila, Amr Mohamed, Mohsen Guizani, And Mounir Hamdi. "TIDCS: A Dynamic Intrusion Detection and Classification System Based Feature Selection". Digital Object Identifier June 3, 2020.
14. Marella, V. C., Veluru, S. R., & Erukude, S. T. (2025, September). FedOnco-Bench: A Reproducible Benchmark for Privacy-Aware Federated Tumor Segmentation with Synthetic CT Data. In 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 870-876). IEEE.
15. T. Kim and W. Pak, "Early Detection of Network Intrusions Using a GAN-Based One-Class Classifier," in IEEE Access, vol. 10, pp. 119357-119367, 2022.
16. Ullah I., Mahmoud Q.H. (2020) A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Nnetworks. In: Goutte C., Zhu X. (eds) Advances in Artificial Intelligence. Canadian AI 2020.
17. Y. K. Al-Douri, V. Pangracious and M. Al-Doori, "Artifical immune system using Genetic Algorithm and decision tree," 2016 International Conference on Bio-engineering for Smart Technologies (BioSMART), Dubai, United Arab Emirates, 2016, pp. 1-4,
18. Al-Sharhan, Salah. (2010). Artificial Immune Systems – Models, Algorithms And Applications. International Journal of Research and Reviews in Applied Sciences. 3.
19. W. A. Nassan, T. Bonny, K. Obaideen and A. A. Hammal, "AN EHDL-IDS model-based Prediction of Chaotic System: Analyzing the Impact of Training Dataset Precision on the Performance," 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 2022, pp. 337-342
20. https://research.unsw.edu.au/projects/unsw-nb15-dataset