International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Enhanced Intrusion Detection System in IOT-MQTT Using Machine Learning Classifiers

Ms.R.Leelavathi, M.E.[1],Dr PMVijayan,M.Tech[2],(ph.D)

V.SaiMounika[3], S.GoldenReddy4, B.Suvarna5, B.Vikas6, N.Tejeshwar[7]

[1]*Associate Professor of ECE dept, Siddartha Institute of Science and Technology JNTUA University, INDIA*

[2]*Associate Professor of ECE dept, Siddartha Institute of Science and Technology JNTUA University, INDIA*

[3,4,5,6,7]*Students of ECE dept, Siddartha Institute of Science and Technology JNTUA University, INDIA*

E-mail:saimounikaveluru@gmail.com

Abstract:

This study significantly enhances IoT security by employing state-of-the-art techniques to detect a range of attacks, including as DoS, brute force, and malformed packets. The hybrid classifier that combines a 1D CNN and fuzzy logic is optimized via the Improved Vulture Starvation-based African Vultures Optimisation Algorithm (IVS-AVOA). The system further enhances attack detection by employing an ensemble classifier that uses models such as Random Forest, XGBoost, and LightGBM, augmented by TSO. The hybrid solution has a 98.5% detection accuracy and a 1.2% false-positive rate, which is better than standalone rule-based orML-driven IDS. The solution offers a robust and adaptable defence mechanism, improving security in vast IoT contexts by utilizing a publically available MQTT dataset and scalable, efficient, and dependable methodologies.

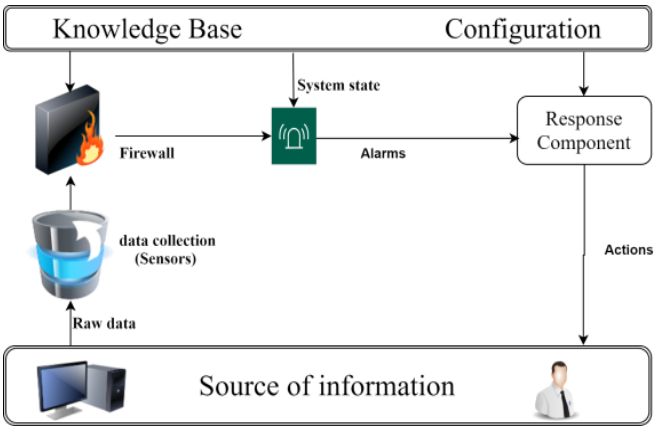
Keywords**:** *IoT Security, MQTT protocol, IDS, Machine learning classifiers (Fuzzy Logic, 1DCNN)*

I. INTRODUCTION

A security tool called an intrusion detection system (IDS) monitors data and network activity to identify any unauthorized or suspicious activities. When it finds irregularities, it notifies users and helps identify potential cyber attacks or unauthorized access. An IDS's primary goal is to increase security by identifying potential threats before they have a chance to do damage. It also collects evidence for further investigation. When combined with firewalls and antivirus software, an IDS offers stronger protection. While firewalls block unwanted traffic, IDS provides detailed analysis and continuous monitoring to detect threats that might get through. Traditional IDS systems are mainly of two types:

1. SIDS, or signature-based intrusion detection system, compares network traffic to a database of recognised attack patterns, or "signatures." If any recorded signatures are matched by the traffic, the system notifies the user. SIDS's inability to recognize unexpected or original attacks is a disadvantage, despite its exceptional ability to detect previous attacks. To remain useful, it must regularly update its signature database. More recent, dynamic cyber attacks could be problematic for SIDS.

2. A technology known as anomaly-based intrusion detection systems, or AIDS, monitors normal network behaviour over time to spot new or unknown threats. AIDS Anomaly-Based IDS: This technique is intended to identify novel or unidentified assaults by tracking typical network activity over time. It sets a baseline for what constitutes "normal" behaviour and sounds an alert when it notices behaviour that doesn't fit this pattern. This implies that it can detect dangers that are not yet known, such zero-day attacks, which take advantage of security holes before they are noticed. AIDS is more adaptable than SIDS, but if the baseline isn't established properly, it might occasionally cause false alarms. Modern intrusion detection systems (IDS) have improved beyond older methods like Signature-Based(SIDS) and Anomaly-Based(AIDS) IDS, using



new techniques to make detection more accurate, reduce false alarms, and better identify cyber attacks. These systems include Hybrid IDS, which combines the strengths of both SIDS and AIDS; Machine Learning-Based IDS, which uses AI to analyse network activity and spot threats; Deep Learning- Based IDS, which is better at detecting complex attacks with fewer. False alerts; Cloud-Based IDS, designed to monitor cloud services and detect issues like DoS attacks or unauthorized access; and Block chain-Based IDS, which uses block chain technology to protect data logs, ensuring they can't be tampered with. These advanced IDS systems work by real time monitoring, using AI, and employing deception technology to prevent and detect intrusions, making them much more effective against today's cyber threats.

II. METHODOLOGY

Intrusions and attacks are any unauthorized actions meant to breach the security of a network or device in order to get, change, or destroy private data, disrupt corporate operations, or take over the system. An incursion or attack is any unapproved activity intended to compromise a network's or device's security in order to gain access to, alter, or destroy private information, interfere with business activities, or take over the system. When hackers take advantage of flaws in IoT devices or communication networks, intrusions are produced. These flaws could result from things like: Intrusions occur when hackers exploit vulnerabilities in communication networks or IoT devices. These problems could result from the following: These flaws could be caused through:

• Insufficient security protocols. Insecure communication protocols allow hackers to intercept or manipulate data, as well as weak passwords and unpatched vulnerabilities.

• Devices with insufficient memory or processing power make it easier for hackers to penetrate the system. Intrusion detection systems (IDS) detect intrusions by continuously monitoring device activity and network traffic. These algorithms look for unusual or problematic tendencies. After detecting an intrusion, the IDS uses anti-virus software. This software scans files and network data to detect known malicious software and malicious actions on IoT devices.

MQTT stands for Message Queuing Telemetry Transport is a lightweight messaging protocol used for communication between devices, especially in IoT systems. It's designed to be efficient and to work well with limited resources, such as devices with low processing power or bandwidth.

MQTT is a widely used protocol for machine-to-machine (M2M) communication due to its simple architecture and ability to adjust to irregular connectivity. Attacks may happen at different times:

- Data interception: Cybercriminals can use private information by intercepting device-to-device conversations.
- In addition to disturbing the functionality of IoT devices by gaining unauthorised access, malicious actors may overload the MQTT broker with traffic, making the system inaccessible to approved users. We call this Denial of Service (DoS).

Several factors make MQTT superior to alternative protocols including CoAP, HTTP, and XMPP. Compact, effective communication, and dependable delivery of messages For intrusion detection, machine learning (ML) and deep learning (DL) are recommended due to their:

• Adaptability: They can adapt the one attack strategies by learning from past data.

• High accuracy: CNNs and other algorithms are excellent at identifying minute patterns and irregularities

• Scalability: ML and DL are perfect for complicated networks because they can process massive amounts of data from IoT devices.

• Automated detection: It eliminates the need for human intervention by detecting and reacting to intrusions in real time.

III. LITERATURE SURVEY

[1] Vaccari, "MQTTset, a New Dataset for Machine Learning Techniques on MQTT," 2023

Introduces the MQTTset dataset for IoT networks, combining cyber-attacks and legitimate data to train machine learning models for attack detection. Sidharthan et al. (2022) developed an intrusion detection system (IDS) using Elite Machine Learning algorithms and a lightweight protocol. They collected

sensor data from three environments, generated multi-context features and evaluated performance using metrics like F1-score, accuracy, and prediction, focusing on attack traffic.

[2] Haripriya and Kulothungan, "Secure MQTT: an efficient fuzzy logic-based approach to detect DoS IN MQTT for IoT," 2019 created the "Secure-MQTT"model using fuzzy logic to detect intrusions and protect low configured IoT devices from DoS attacks. Bedi et al. (2021) applied machine learning algorithms (ANN, Logistic Regression, Random Forest, SVM, Decision Tree) to detect anomalies in IoT devices, evaluating performance with metrics like precision, recall, accuracy, and F1-score. Provide the MQTTset dataset, which combines real data with cyber attacks to train machine learning models for attack detection in IoT networks. In order to create an intrusion detection system (IDS), Siddharthan et al. (2022) used Elite Machine Learning techniques with a lightweight protocol.

[3] In their article "Implementing optimized classifier for attack distribution& BAIT based attack correction in IoT," Babu and Veena (2021) suggested an attack detection system for IoT that uses a standard dataset to compute Euclidean distances in order to identify the shortest data transfer channels. They used an Optimized Deep Belief Network (DBN) and the Whale with Distance-based Update (W-DU) algorithm to optimize weight. The system's effectiveness was compared to more conventional models. The Metaheuristic was employed by Krishna and Thangavelu (2021) for intrusion detection and data pre-processing.

[4] Li et al, "ADRIoT", (2021) developed the "ADRIoT" anomaly detection method for Internet of Things networks, which adds a "multi-edge collaborative mechanism "to boost load capacity while reducing edge processing. By successfully identifying a range of IoT-based dangers, our technique contributed to the development of a more secure IoT infrastructure. The improved intrusion detection system (IDS) presented by Bhosale and Sonavane (2021) uses Received Signal Strength Indicator (RSSI) data to identify attacker nodes and their neighbours in order to detect wormhole assaults in IoT devices. Due to the system's excellent energy economy, it can be used in IoT situations with limited resources.

IV. RESEARCH GAP

Existing models has several limitations, including the inability to tune hyper-parameters in any machine learning approaches and the lack of multi-model traffic classifiers. It suffers from a low convergence rate, which slows down the process in local optima, and loses important data during processing, creating challenges. Its poor convergence rate causes problems by slowing down the process in local optima and losing crucial data while processing. Memory and data rate are adversely affected by the vast number of factors. Its poor exploitation rate and inability to use a large number of classifiers or datasets result in a more complex system. The model's issues with time complexity and attack stage analysis lead to over fitting and computing complexity. Additionally, it doesn't use optimization techniques, doesn't mention processing time, ignores negative measures, performs worse as the number of hops rises, and doesn't identify simultaneous attacks.

V. PROPOSED WORK

In the proposed method we introduce a hybrid IDS for IoT devices that combines a machine learning classifier with an optimization method to detect intrusions in IoT networks. First, the MQTT dataset is used, and then preprocessing— which includes data normalisation and cleaning comes next. Three feature sets are chosen following preprocessing. Weighted fused features are then produced by combining these sets using a weight parameter. IVS-AVOA is used to optimally choose the first set of features, IVS-AVOA is used to optimise the second set, which is generated from statistical features, and IVS-AVOA is used to optimise the third set, which is produced by an auto encoder. An improved AVOA technique is then utilised to adjust the hyper parameters of a Hybrid Classifier (HC), which incorporates fuzzy logic and 1D-CNN, using the fused features as input. By averaging the data, the final categorized outcome is produced, improving intrusion detection accuracy.The MQTT dataset is used to detect different types of attacks, such as brute force authentication, malformed data, SlowITe, MQTT publish flood, and DoS flooding. In the final step, the weighted features are analysed in an HC-based IDS phase, where fuzzy logic and IDCNN detect intrusions. Parameters like the exponent bound in fuzzy logic and filters in IDCNN

are optimized with IVS-AVOA to improve detection accuracy". The weighted features are then examined in an HC-based IDS phase, where intrusions are found using fuzzy logic and IDCNN. To increase detection accuracy, IVS-AVOA is used to optimise parameters such as the filters in IDCNN and the exponent bound in fuzzy logic. The goal of this project is to develop an intrusion detection system (IDS) architecture that ensures secure data transit between nodes while detecting intrusions in IoT devices using hybrid classifiers and heuristic parameteroptimisation.Anew intrusion detection method called HC is introduced, which combine fuzzy logic and 1DCNN. The IVS-AVOA was developed to modify parameters in both fuzzy logic and 1DCNN for more ultimately improving the accuracy of intrusion detection in the network. The flow diagram below outlines the method in detail for identifying intrusions or assaults in the MQTT protocol used by the Internet of Things. Let's talk about each block in

[1] MQTT Dataset

The MQTT dataset, based on IoTFlock, simulates network traffic for IoT devices using CoAP and MQTT protocols. It includes cyber threats such as publish floods, packet targeting IoT networks. In order to improve the network's intrusion detection accuracy, the IVS-AVOA method is also utilised to choose the best features, optimise weights during feature fusion, and modify the membership function in fuzzy logic and filters in 1DCNN.

To test detection methods, the MQTTset contains actual cyber attacks that target the MQTT network. The MQTTset include threats such as flooding DoS, SlowITe, MQTT publish flood, malformed data (sending corrupted packets), and brute force authentication. The MQTT dataset is produced by identifying both benign and malevolent activity by extracting important information from unprocessed network traffic. IP addresses, ports, communication durations, MQTT client information, and other network elements are eliminated as unnecessary information. Effective intrusion detection is then carried out using the filtered data.

[2] DATA PRE PROCESSING:

MQTT data is cleaned during the preparation phase to ensure accuracy by removing duplicates and correcting errors. Data normalization is the used to scale the data eliminate noise. Feature extraction is done using the preprocessed data that is produced.

[3] FEATURE SELECTION: Three models are used in the feature selection procedure. The optimal initial set of features is chosen directly from the pre-processed data in Model 1 using the IVS-AVOA. The IVS-AVOA is used to select the best second set of features after statistical features are taken out of the preprocessed data in Model 2. In Model 3, deeper features for intrusion detection are extracted from the pre-processed data using an auto encoder, and the best third party is chosen using the IVS-AVOA.

[4] WEIGHT OPTIMIZATION: The weighted feature selection step of the IoT intrusion detection model makes use of the three sets of chosen features (F1c, F2c, and F3c). First, the optimum weights (W1 and W2) identified by IVS-AVOA are used to pick the most significant characteristics from Feature Sets 1and 2. They are combined with Feature Set 3 to produce weighted fused features. with the optimized weight (W3) also calculated using IVS- AVOA. The optimal weight (W3) is also obtained using IVS-AVOA. Increasing the accuracy of intrusion detection is the goal of the weighted feature fusion stage.

$$\text{Accuracy (AR)} = \frac{TP+TN}{TP+TN+FP+FN}$$

[5] PROPOSED IVS-AVOA OPTIMIZATION

The AVOA (African Vulture Optimization Algorithm) was developed with inspiration from the way African vultures search for food. It follows these four fundamental steps: Initially, it weighs every option to find the greatest one, which is referred to as the "optimal vulture. "Next, using certain control parameters (β = 0.2, R1 and R3 = 0.6), the novel method automatically adapts to changing conditions to compute the vulture hunger rate. Following that, the algorithm randomly investigates new alternatives when there is a significant discrepancy between the optimal solution and the present solution. Ultimately, it shifts to exploitation, honing and concentrating on the optimal option when the gap is less. To put it briefly, the algorithm simulates vultures alternating between scouting new territory and honing what they have already discovered.

PROPOSED METHODOLOGY DIAGRAM

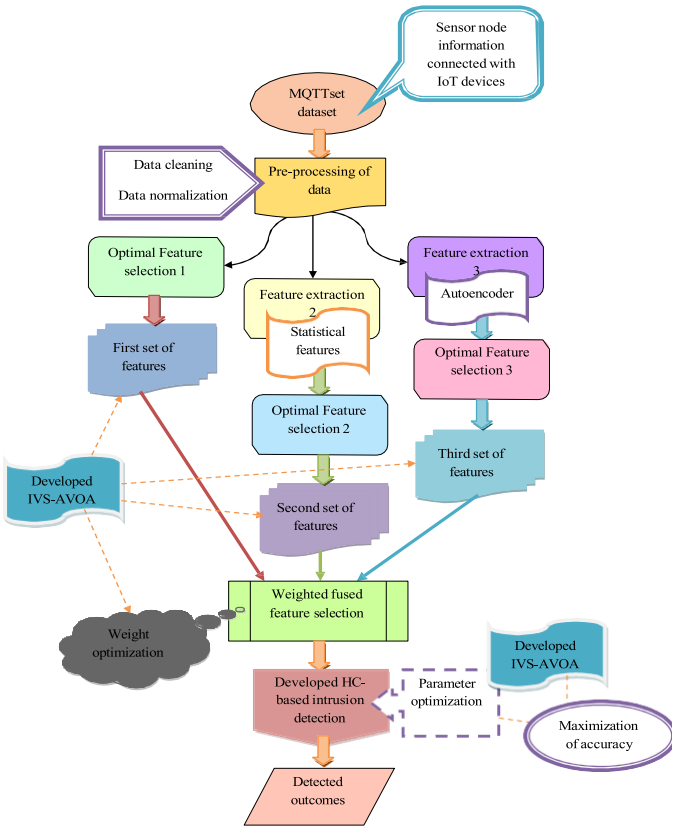


Figure: Combination of fuzzy and 1DCNN using the developed IVS-AVOA

[6] PROPOSED HYBRID CLASSIFIER

Fuzzy and 1DCNN Hybrid Classifier In order to increase intrusion detection accuracy, the IoT-based IDS framework uses a 1DCNN in conjunction with fuzzy logic. Although fuzzy logic enhances detection performance, the 1D-CNN has benefits like improved classification, quicker processing, and the ability to automatically learn features.

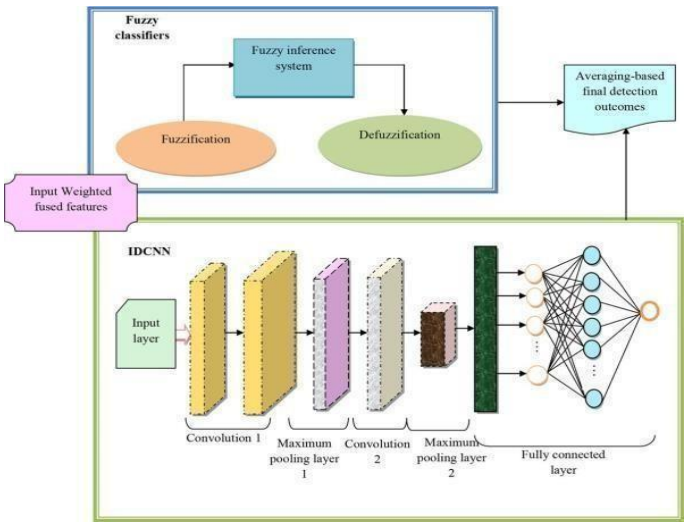


Fig: Developed HC-Based Intrusion Detection System

The IVS-AVOA optimization technique improves the accuracy even further by fine-tuning the 1D-CNN filters and altering the exponential bound in fuzzylogic. By fine-tuning the fuzzy exponential bound between 0.01 and 0.09 and optimizing the 1D-CNN filters between 64 and 128, a more accurate intrusion detection system is produced. The developed IDS for IoT devices are validated with several quantitative measures such as "specificity, sensitivity, MCC, F1-score, FNR, FPR, NPV, precision".

VI. RESULTS AND DISCUSSION

Python was used as the programming platform for the implementation of the suggested intrusion detection system framework for the Internet of Things. The accuracy performance and F1-score of the five distinct heuristic-based optimization techniques at various learning percentages are shown in figures a and b below. The new model, IVS-AVOA-HC, was compared to machine learning techniques such as Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), Fuzzy Logic, 1D Convolutional Neural Networks (1DCNN), and Fuzzy-1DCNN, as well as a number of heuristic approaches, such as Particle Swarm Optimization (PSO)-HC, Jaya Algorithm (JA)-HC, Brainstorm Optimisation (BSO)-HC, and AVOA-HC. As the learning percentage rises, all approaches show a growing trend in accuracy and F1score, as seen in the figure. This suggests that improved model performance is correlated with more data.

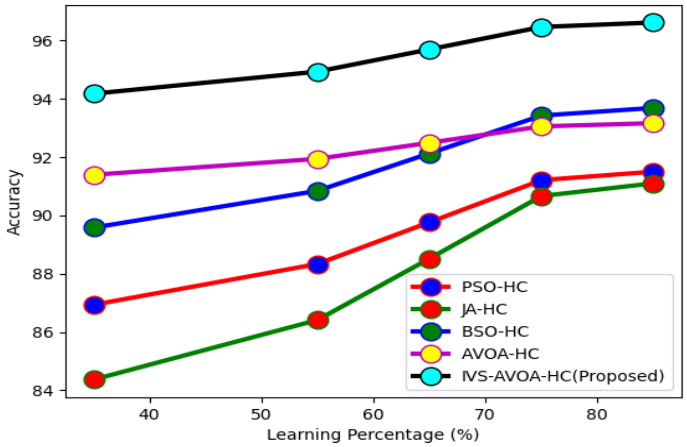
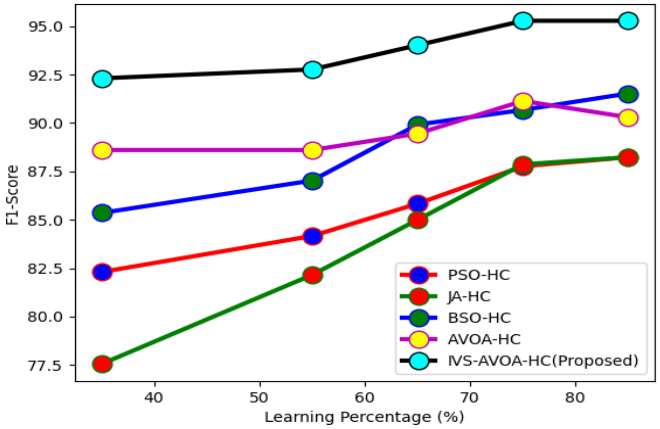


Figure a: Accuracy

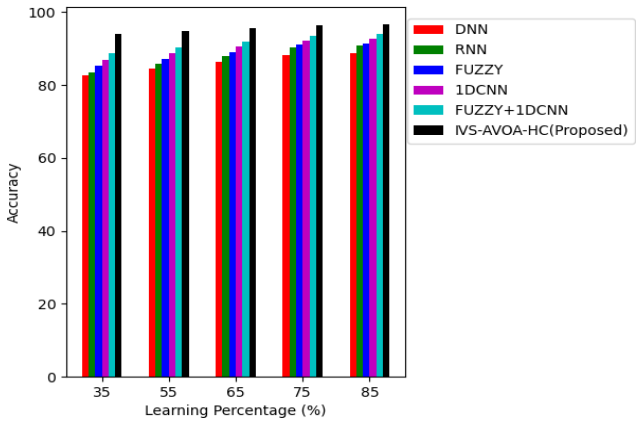
The above figure a and below figure b shown the proposed IDS framework for IoT devices with meta-heuristic algorithms over "(a) accuracy, (b) F1- score. The below following "figures c and d" presents the

detection performance of the proposed IDS framework for IoT devices is evaluated through analysis with different classifiers through analysis with different classifiers. When comparing the conventional and Proposed technique, it is confirmed that the proposed technique (Fuzzy-1DCNN) has highest "Accuracy and F1-scale" than the **DNN, Fuzzy, RNN, 1DCNN**. So, the efficiency of the proposed model is high than the conventional techniques. Similarly, upon examining below table a; the F1-score, MCC& FPR and precision values of the proposed method. Classifiers like **DNN, Fuzzy-1DCNN, RNN, and Fuzzy,** respectively", improved accuracy performance compared to the other Hybrid classifiers.



Figureb:F1-Score

The following figures c and d presents the detection performance of the proposed IDS framework for IoT devices is evaluated through analysis with different classifiers through analysis with different classifiers. When comparing the conventional and proposed technique.



it is confirmed that the proposed technique (Fuzzy1DCNN) has highest "Accuracy and F1-scale" than the DNN, Fuzzy, RNN, 1DCNN . So efficiency of the proposed model is high than the conventional techniques.

Measures	DNN	RNN	Fuzzy	1DCNN	Fuzzy-1DCNN	IVS AVOA-HC
Accuracy	88.73	90.72	91.45	92.64	93.95	96.63
F1-score	84.54	87.37	88.32	90.26	92.74	95.27
MCC	75.51	80.18	81.69	84.72	88.62	92.62
NPV	88.76	90.75	91.42	92.65	93.52	96.68
Precision	81.48	84.52	84.86	87.74	90.82	94.06

Table a : comparative validation of the proposed IDS framework for IoT devices with existing classifiers

"The above tabulation shows a comparative validation of the proposed IDS framework for IoT devices with existing classifiers. From the table above, it is evident that the proposed IVS-AVOA-HC achieves the highest accuracy, F1-score, MCC, NPV, and with values of 96.63%, 95.27%, 92.62%, 96.68%, and 94.06%, respectively". As seen in the tabular column, the proposed IVS- AVOAHC has improved accuracy by 7.9% compared to DNNHC. Similarly, the proposed framework shows accuracy improvements of 5.91%, 4.88%, 3.99%, and 2.68% over the Hybrid Classifiers like RNN, DNN, Fuzzy, 1DCNN, Fuzzy-1DCNN respectively".

Measures	PSO-HC	JA-HC	BSO-HC	AVOA-HC	IVS-AVOA-HC
Accuracy	91.51	91.12	93.63	93.15	96.65
F1-score	88.61	88.55	92.34	93.21	95.22
MCC	82.07	81.96	87.95	81.85	88.10
NPV	91.56	91.05	93.67	93.15	96.61
Precision	86.04	86.65	90.71	90.24	95.32

Table b : Comparative Validation of the Proposed IDS Framework for IoT Devices With Existing Meta-heuristic algorithms

Similarly on examining the above tabular column b; the accuracy, F1-score, MCC, NPV and Precision valuesoftheproposedmethod. The above table shown the comparative validation of the proposed IDS framework for IoT devices with existing meta-heuristic algorithms. As seen in the tabular column, the proposed IVS-AVOA-HC has improved accuracy by compared to PSO-HC, JA-HC, BSO-HC, AVOA -HC, I V S -AVOA-HC, respectively.

VII. CONCLUSION

In this work, we created an Intrusion Detection System (IDS) for the Internet of Things (IoT) that improves intrusion detection performance by combining hybrid classifiers with an optimization approach. To guarantee accurate results, data pretreatment techniques including cleaning and normalization were performed on the assessment dataset, which was gathered from MQTT. Three sets of optimal features were extracted: one from statistical features, one from auto encoder-based deep features, and one from the pre-processed data. During the feature fusion phase, these ideal features were fused after being chosen using the IVS-AVOA optimization approach.

Fuzzy logic and 1DCNN were utilized to categorize intrusions during the intrusion detection phase. During the intrusion detection stage, intrusions were classified using 1DCNN and fuzzy logic. Compared to DNN, RNN, Fuzzy, 1DCNN, and Fuzzy-1DCNN, the suggested IVS-AVOAHC-based IDS performed better than conventional classifiers, with accuracy gains of 8.89%,6.44%,5.67%,4.24%,and2.83%,
respectively. We deduce from these findings that the suggested IDS framework provides noticeably superior performance and efficiency when compared to traditional techniques for IoT intrusion detection. In Future, we developed our project “Enhanced Intrusion Detection System in IoT-MQTT using Machine Learning classifiers” (Fuzzy and 1DCNN). Future work could incorporate Hybrid classifiers Deep-Learning models Like CNN+LSTM Transformer based models, or even attention mechanisms tohandle complex attack patterns and improve detection accuracy.

VIII. REFERENCES

[1] A. Fatani, M. Abd Elaziz, A. Dahou, M. A. A. AlQaness and S. Lu, "IoT Intrusion Detection System Using Deep Learning and Enhanced Transient Search Optimization," IEEE Access, vol. 9, pp. 123448-123464, 2021.

[2] D.Breitenbacher,I.Homoliak,Y. L.Aung,Y. Elovici and N. O. Tippenhauer, "HADES-IoT: A Practical and Effective Host-Based Anomaly Detection System for IoT Devices (Extended Version)," IEEE Internet of ThingsJournal,vol.9,no.12,pp.9640-9658,15June 15, 2022.

[3] Abdollahzadeh, B., F. S. Gharehchopogh, and S. Mirjalili. 2021. African vulture’s optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems.

[4] Bedi, P., S. Mewada, R. A. Vatti, C. Singh, K. S. Dhindsa, M. Ponnusamy, and R. Sikarwar. 2021. Detection of attacks in IoT sensors networks using machine learning algorithm. Microprocessors and Microsystems82:103814.

doi:10.1016/j.micpro.2020.103814.

[5] Ciklabakkal, E.,A.Donmez, M.Erdemir,E.Suren, M. K. Yilmaz, and P. Angin. 2019. ARTEMIS: An intrusion detection system for MQTT attacks in internet of things. 2019 38th Symposium on Reliable Distributed Systems(SRDS).
doi:10.1109/SRDS47363.2019.00053.

[6] Liu, J., D. Yang, M. Lian, and M. Li. 2021. Research on intrusion detection based on particle swarm optimization in IoT. IEEE Access 9:38254–68. doi:10.1109/ACCESS.2021. 3063671.