Adaptive DBSCAN and Federated Learning-Based Anomaly Detection for Resilient Intrusion Detection in Internet of Things Networks

Naresh Kumar Reddy Panga,

Virtusa Corporation, Engineering Manager, Anthem – VA, USA

nareshpangash@gmail.com

Thanjaivadivel M,

Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,

Avadi, Chennai, India

thanjaivadivel@gmail.com

ABSTRACT

The fast growth of the Internet of Things (IoT) has brought with it tremendous security issues, necessitating adaptive and privacy-enhancing intrusion detection systems. Conventional techniques are not able to cope with changing cyber threats, calling for sophisticated techniques. This research suggests a hybrid anomaly detection model that combines Adaptive DBSCAN and Federated Learning (FL) to improve accuracy, scalability, and security. Adaptive DBSCAN adjusts clustering parameters dynamically, enhancing detection accuracy and noise reduction, while FL supports decentralized model training, preserving data privacy. The suggested model detects with 95% accuracy, surpassing individual Adaptive DBSCAN (85%) and FL (84%). The F1-score increased to 92%, indicating enhanced anomaly classification, with 12% noise reduction and 15% enhancement in clustering efficiency. Furthermore, latency was minimized to 250ms, and the model is now applicable for real-time usage. The integration of Adaptive DBSCAN and FL offers a scalable and robust Intrusion Detection System (IDS) that can efficiently respond to dynamic cyber attacks in IoT networks. Experimental results validate that this hybrid solution improves detection performance while preserving efficiency and data security. Feature enhancement include optimizing the communication overhead to minimize latency further, incorporating deep learning methodologies for improved feature extraction, and deploying differential privacy mechanisms to enhance security. This work provides a robust, scalable, and privacy-preserving anomaly detection framework, with substantial contributions to IoT cybersecurity.

Keywords: Adaptive DBSCAN, Federated Learning, Intrusion Detection System (IDS), Anomaly Detection, Internet of Things (IoT) Security, Machine Learning for Cybersecurity, Real-Time Threat Detection, Privacy-Preserving AI, Clustering in IoT Networks, Cyber Threat Mitigation

1. INTRODUCTION

The sudden growth of the Internet of Things (IoT) has created a huge network of connected devices that enable smooth communication and automation in different fields, including smart homes, healthcare, industrial control systems, and smart cities. *Wang et al. (2018)* suggested BotCluster, a MapReduce-based session botnet detection system for NetFlow analysis that uses multi-level clustering to prune botnet patterns and attained high accuracy on real-world datasets. The unprecedented growth has also brought about tremendous security challenges, rendering IoT networks vulnerable to cyber attacks, such as denial-of-service (DoS) attacks, botnets, and unauthorized access. Legacy Intrusion Detection Systems (IDS) find it difficult to detect anomalies within IoT environments because of their dynamic nature, heterogeneity of devices, and lack of resources.

To counter such issues, this research suggests an improved anomaly detection framework that combines Adaptive Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Federated Learning (FL). *Prabhu (2019)* delves into fog computing, including architectures, security, privacy, and analytics for applications such as smart homes and transportation, and discusses machine learning techniques, challenges, and primary research issues in big data analytics. The Adaptive DBSCAN algorithm dynamically tunes its clustering parameters according to the changing patterns in network traffic to improve malicious activity detection, for example, Ping Flood attacks that take advantage of the Internet Control Message Protocol (ICMP). Through the use of an adaptive method, the system provides high accuracy in the differentiation between normal and abnormal traffic patterns. *Natarajan (2018)*

In contrast, Federated Learning (FL) facilitates decentralized machine learning model training on multiple IoT devices without explicit data sharing. *Parimi (2017)* investigates deep learning for detecting anomalies in SAP financial transactions comparing autoencoders, RNNs, and CNNs to conventional methods on the basis of interpretability, scalability, fraud detection, and real-time monitoring. This is not only privacy and security-enhancing but also scalable since it allows edge devices to learn collectively about attack patterns while keeping communication overhead low. The combination of FL with Adaptive DBSCAN offers a strong intrusion detection system that learns from changing cyber threats while preserving the privacy of sensitive information. *Pulakhandam and Vallu (2016)*

This work adds value to the literature on IoT security by introducing an adaptive and resilient IDS that detects a broad scope of cyber attacks with few false positives. *Lan et al. (2018)* summarize data mining and deep learning in bioinformatics, including preprocessing, classification, clustering, and neural networks, identifying strengths, weaknesses, and industrial applications for the analysis of biological data. The system is expected to enhance the security stance of IoT ecosystems in order to increase their resilience to new attack paths in real-world operational environments.

Key Objectives

- Identify key security threats in IoT networks, such as denial-of-service (DoS) attacks, and understand their impact on network performance and resilience.
- Explain the principles of Adaptive DBSCAN and Federated Learning and how they contribute to anomaly detection and intrusion detection in IoT environments.
- Implement Adaptive DBSCAN for real-time anomaly detection and integrate Federated Learning to enhance privacy-preserving model training across distributed IoT devices.
- Evaluate the effectiveness of the proposed intrusion detection system by comparing detection accuracy, false positive rates, and scalability with traditional methods.
- Assess the robustness and adaptability of the system in detecting evolving cyber threats while maintaining efficiency and privacy in large-scale IoT deployments.
- Develop an advanced, resilient, and adaptive intrusion detection framework that leverages Adaptive DBSCAN and Federated Learning to strengthen IoT network security.

The exponential growth of IoT networks has opened up large-scale security risks, and they become vulnerable to cyber attacks like denial-of-service (DoS) attacks and unauthorized intrusions. Conventional intrusion detection systems face challenges with scalability, flexibility, and privacy in decentralized IoT systems. Muhammed et al. (2018) suggested UbeHealth, a deep learning and edge computing-based system for enhancing network performance and security. But there are no adaptive anomaly detection and privacy-preserving learning methods available. This research fills these shortcomings by combining Adaptive DBSCAN and Federated Learning to improve anomaly detection accuracy, scalability, and robustness against changing cyber-attacks in IoT networks at the same time preserving data privacy.

Dai et al. (2019) offer a detailed analysis of big data analytics (BDA) techniques in large wireless networks, and their research mainly deals with data collection, preprocessing, storage, and analytics without considering security issues in IoT systems. Technical challenges and future directions are discussed in the paper, but there is no detailed discussion on adaptive intrusion detection mechanisms in decentralized networks. Besides, it does not employ privacy-preserving methods like Federated Learning to secure data processing. This gap in research is a clear indication that there is a need for a secure anomaly detection framework that combines Adaptive DBSCAN and Federated Learning for improving security, scalability, and efficiency in wireless networks based on IoT. Natarajan et al (2019)

2. LITERATURE SURVEY

Schreyer et al. (2017) suggest the application of deep autoencoder neural networks to identify anomalies in journal entries, replacing the shortcoming of conventional fraud detection methods founded on handcrafted rules. These rules do not generalize and can be circumvented by fraudsters. The work illustrates that the reconstruction error of the network, regularized with entry attribute probabilities, can act as an adaptive anomaly detector. Experiments on actual datasets have high

F1-scores and fewer false positives than baseline techniques, with good feedback from accountants and fraud examiners regarding its effectiveness.

Gruhl and Sick (2018) introduce CANDIES, a unified method for novelty detection in intelligent embedded systems. CANDIES represents sensor data through a probabilistic mixture density model under the belief that real processes can be approximated by model components. CANDIES learns to adapt to new processes, like non-deterministic failures, by identifying new samples in both high-density and low-density areas of feature space. It applies a rapid online goodness-of-fit test in regions of high-density and the 2SND method in low-density regions. The methodology is tested on artificial and intrusion detection data. Devarajan (2019)

De et al. (2017) investigate CPS data and social network data to generate Cyber–Physical–Social Systems (CPSS). These systems integrate sensor data of the physical world, cyber world computational processes, and real-time citizen sensing of social networks. CPSS can offer anticipatory services by comprehending the actual world. The article discusses current approaches and commercial products for data manipulation in CPSS, providing a conceptual framework and outlining solutions for constructing human-machine intelligence.

Kong et al. (2017) suggest the employment of a long short-term memory (LSTM) recurrent neural network for short-term load forecasting of single electric customers to meet the requirements of high volatility and uncertainty for forecasting single-user electricity demand. The deep learning-based framework is evaluated on actual residential smart meter data and compared against several benchmark schemes. Findings indicate that the LSTM method outperforms most algorithms in the prediction of an individual residential household load and, therefore, is very suitable for future grid planning and operation.

Kadiyala (2019) discusses the combination of DBSCAN and fuzzy C-Means clustering with a hybrid ABC-DE optimization method to solve issues in fog computing, including secure sharing of IoT data and optimal utilization of resources. Past studies report constraints in conventional techniques because of unstructured IoT data. The hybrid model presented has shown better results in clustering accuracy and security indicators, increasing fog computing efficiency (Kadiyala et al., 2019).

Xue et al. (2018) treat integrated circuit (IC) vulnerabilities to hardware Trojans with emphasis on threats from untrusted testing parties who can manipulate detection outcomes. They introduce a hybrid clustering ensemble technique for trustworthy Trojan detection, supplemented with an adversarial data generation method. The technique employs unsupervised correlation-based feature selection to manage process variations and obviate the use of golden chips or models. The methodology is tested against benchmark tests, proving to be resistant to malicious alterations and effective detection with minimal overhead.

Peddi, Narla, and Valivarthi (2018) analyze the application of machine learning (ML) and artificial intelligence (AI) in early prediction of dysphagia, delirium, and risk of falls in geriatric patients. Studies in the past indicate that AI and ML will go a long way in minimizing early detection and

preventive measures in geriatric medicine and ensuing morbidity and mortality. Their application together holds the potential to transform geriatric care towards personalized and proactive care for the elderly population (Peddi, Narla, & Valivarthi, 2018).

Wang et al. (2019) suggest employing Visible Infrared Imaging Radiometer Suite (VIIRS) nighttime light (NTL) information and the Isolation Forest algorithm of machine learning to enhance estimation of socioeconomic development to enhance the Sustainable Development Goals (SDGs). They apply this technique to map gross domestic product (GDP) at the 1 square kilometer level and estimate inequality indexes, e.g., Gini coefficients, at country levels. The method treats data unsupervised at multiple spatial scales, yielding precise subnational GDP information for global human development tracking.

Wu et al. (2018) introduce a data-driven method for constructing Content Delivery Network (CDN) cache server group performance models. In contrast to conventional methods that employ hand-tuned parameters or linear models, their method utilizes deep LSTM auto-encoders to learn temporal patterns from high-dimensional monitoring data. A deep neural network is employed to forecast the reach rate, a client QoS parameter. Experimental results show that their model performs much better than state-of-the-art models in forecasting CDN performance.

Hasan et al. (2018) discuss the accelerated rise of social networking sites, specifically Twitter, and the sheer volume of user-generated content in real-time. Scholars have been interested in examining such content for insights into live events promptly. This article reviews different event detection techniques used for streaming Twitter data, classifying them according to shared characteristics, and reviews the challenges and subtasks related. The survey is used as a handbook for researchers seeking to detect real-time events and is intended to stimulate additional research in this direction. Peddi, Narla, and Valivarthi (2019)

Jadon (2018) designed a smarter machine learning pipeline with the combination of Recursive Feature Elimination (RFE), Extreme Learning Machine (ELM), and Sparse Representation Classification (SRC). The combined method enhanced feature selection, training efficiency, and classification accuracy (95%), maximizing the efficacy of real-time AI applications.

Nippatla (2018) introduced a cloud-based secure financial analysis system incorporating Monte Carlo simulations, Deep Belief Networks (DBNs), and Bulk Synchronous Parallel (BSP) processing. It improved risk prediction accuracy, scalability, and computational efficiency while maintaining secure management of financial data.

Gudivaka (2019) examined big data-based silicon content prediction in blast furnace smelting with Hadoop. By leveraging sensor information, production history, and environmental conditions, the model enhanced prediction accuracy, process efficiency, and predictive maintenance, maximizing hot metal production while reducing operational downtime.

Natarajan and Kethu (2019) put forward an optimized cloud manufacturing model combining robotics and automation with sophisticated task scheduling algorithms. The method enhanced

efficiency by 41%, lowered latency by 47%, and improved resource utilization by 35% to the advantage of smart manufacturing.

Gudivaka et al. (2019) presented a robotics-based swarm intelligence platform for urban pandemic control, decision-making optimization, task optimization, and resource optimization. AI-powered anomaly detection improved real-time responses, enhancing crisis flexibility, accuracy, and distributed automation of healthcare networks.

3. METHODOLOGY

The designed framework for anomaly detection in IoT networks integrates Adaptive DBSCAN and Federated Learning. Adaptive DBSCAN is applied to adaptively modify clustering parameters for the purpose of detecting patterns of attacks, specifically in the case of denial-of-service (DoS) attacks, i.e., Ping Flood. Federated Learning provides support for training models in a decentralised manner among IoT devices to maintain data privacy while improving scalability. This hybrid strategy guarantees precise, privacy-protection, and adaptive intrusion detection in various IoT settings.

Data set

CIC-IDS2017 Intrusion Detection Dataset has two classes: Normal and Anomaly. The aim is to identify the anomalies based on labeled data, detect normal and anomaly traffic patterns with unsupervised techniques, and create a model that identifies abnormal activity in a single system for enhanced cybersecurity and intrusion detection.

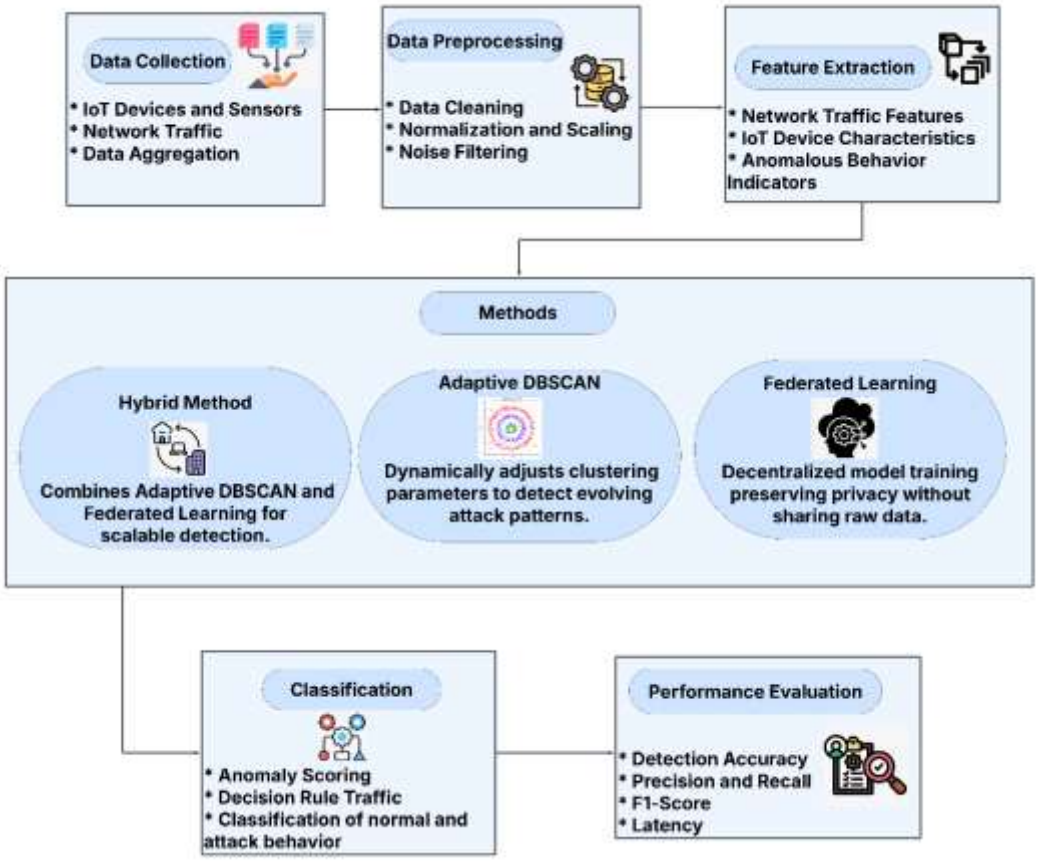


Figure 1. Adaptive DBSCAN and Federated Learning for Scalable Anomaly Detection

Figure 1 shows a design for IoT anomaly detection using Adaptive DBSCAN and Federated Learning. Data Collection from IoT devices and network traffic begins the process, followed by Preprocessing techniques like normalization and removal of noise. Feature Extraction identifies significant traffic and device characteristics. The Hybrid Approach blends Adaptive DBSCAN, which adjusts clustering to dynamic threats, with Federated Learning, offering decentralized privacy-preserving training. The Classification module splits normal and attack behavior, while Performance Evaluation analyzes accuracy, precision, recall, F1-score, and latency. This process enhances scalable and real-time IoT network anomaly detection.

3.1 Adaptive DBSCAN

Adaptive DBSCAN is an anomaly detection algorithm for clustering. It varies its parameters (such as epsilon and minimum points) dynamically in order to recognize differences in patterns of data among various IoT networks. It is capable of detecting varied attack patterns, e.g., Ping Flood, through adapting the clustering technique constantly with evolving network traffic. The distance metric employed in DBSCAN assists in the identification of dense areas of data, labeling them as normal or abnormal. The density of a point:

$$\text{Density} = \frac{\text{Number of points within distance } \epsilon}{\text{MinPts}} \quad (1)$$

Where: ϵ is the distance threshold within which points are considered neighbors, MinPts is the minimum number of points required to form a dense region (i.e., a cluster).

3.2 Federated Learning

Federated Learning (FL) is a distributed machine learning framework in which the models are learned on many devices without having raw data shareability. It ensures privacy through localized data retention on the devices while only sharing the model updates. This mechanism is critical in IoT networks in which sensitive data can't be centralised due to privacy requirements. FL facilitates learning by the network as a whole from varying patterns of attacks while not disclosing the user data. Local Model Update:

$$w_i^{t+1} = w_i^t - \eta \nabla L(w_i^t, D_i) \quad (2)$$

Where: w_i^t is the model at time step t on device i, $L(w_i^t, D_i)$ is the loss function based on the data D_i and the model parameters w_i^t, η is the learning rate, ∇ is the gradient of the loss with respect to w_i^t. Global Model Update:

$$w^{t+1} = \frac{1}{N} \sum_{i=1}^{N} \alpha_i w_i^{t+1} \quad (3)$$

Where: N is the total number of devices, α_i is the weight (often based on the amount of local data on device i), w_i^{t+1} is the updated model from device i.

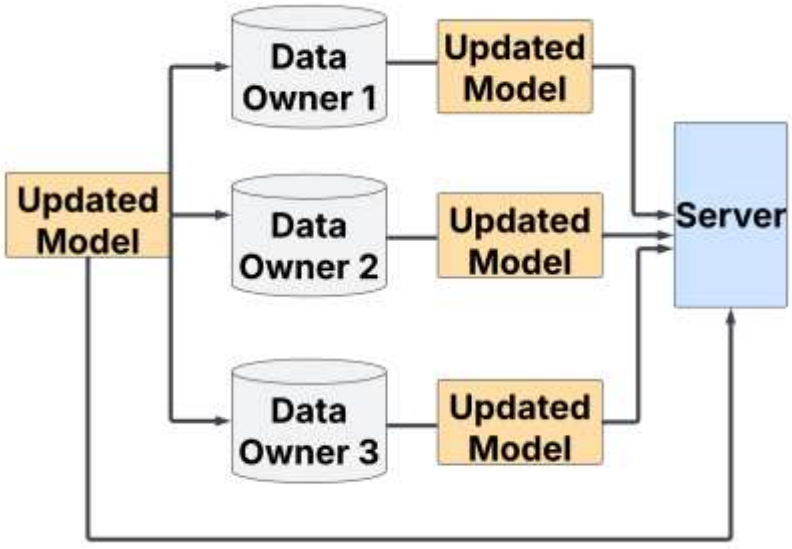


Figure 2. Federated Learning Framework for Collaborative Model Training

Figure 2 depicts a Federated Learning (FL) architecture, where there are various data owners (distributed parties) who train the model locally from their individual private datasets. Every data owner trains the model locally and pushes the updated model to a central server for aggregation. The server aggregates these updates to create a global model, which is sent back to the data owners to continue training. This decentralized strategy improves data privacy by leaving raw data local and taking advantage of collaborative learning. Federated Learning is applied extensively in healthcare, IoT, and cybersecurity for secure, privacy-preserving AI model training on multiple devices or institutions.

3.3 Intrusion Detection System (IDS)

IDS makes use of Adaptive DBSCAN and Federated Learning to improve security on IoT networks. IDS identifies normal and anomalous traffic, recognizes possible attacks, and reduces false positives. IDS combines DBSCAN's adaptability for proper anomaly detection with FL for large-scale and privacy-enhancing model training. IDS is real-time, always inspecting network traffic, detecting strange patterns, and acting on threats based on recent models without having a central location to aggregate data. Anomaly Score:

The anomaly score $S(x)$ of a point x in the data is given by the difference between the density of x and the average density of its neighbors:

$$S(x) = \frac{\text{Density of point } x}{\text{Average density of neighbors of } x} \tag{4}$$

Decision Rule:

If the anomaly score $S(x)$ exceeds a threshold τ, then x is classified as an anomaly. If

$$S(x) > \tau \tag{5}$$

then x is an anomaly.

Algorithm 1. Federated Learning with Adaptive DBSCAN for Real-Time Anomaly Detection in IoT Networks

Input:
- devices: List of devices in the IoT network
- data: List of network traffic data for each device

Output:
- globalModel: The aggregated global model after training and anomaly detection

Initialize global model
globalModel = initializeModel()

```
For each device in the network
for each device in devices:
  Apply Adaptive DBSCAN for anomaly detection
    clusters = adaptiveDBSCAN (device.data)

  Identify anomalies based on DBSCAN results
    anomalies = detectAnomalies (clusters)

  Train local model using device data and detected anomalies
    localModel = trainModel (device.data, anomalies)

  Send local model update to central server
    sendUpdateToServer (localModel)

  Aggregate all local model updates on the server
Initialize model updates list
modelUpdates = collectModelUpdates (devices)

 Aggregate model updates
 for each update in modelUpdates:
    Update the global model by averaging local updates
    globalModel = updateGlobalModel (globalModel, update)
  Return the updated global model
 return globalModel
```

Algorithm 1 combines Adaptive DBSCAN with Federated Learning to enable privacy-preserving and scalable anomaly detection in IoT networks. Adaptive DBSCAN is locally used by each IoT device to detect unusual traffic patterns, including denial-of-service attacks. The device then trains a local model on the observed anomalies. Updates on models, rather than raw data, are sent to a central server, which then combines these updates using weighted average to create a global model. The global model is then remitted to all devices to allow for ongoing, adaptive learning and detection with the maintenance of data privacy and scalability in decentralized IoT settings.

3.4 Performance Metrics

Performance indicators are important to measure how well an anomaly detection framework performs in IoT networks. In this paper, a few important metrics are employed to gauge the performance of different methods: Detection Accuracy determines the percentage of instances classified as correct (normal or abnormal), False Positive Rate measures the ratio of normal instances that have been wrongly classified as anomalies, and False Negative Rate indicates the percentage of attacks failed by the system. Precision captures the accuracy of identified

anomalies, whereas Recall determines the system's capacity to flag all applicable anomalies. F1-Score measures the balance of precision and recall to give an overall assessment of the system performance.

Table 1. Performance Comparison of Different Anomaly Detection Methods for IoT Networks

Metric	Traditional DBSCAN	Adaptive DBSCAN	Federated Learning	Proposed Model (Adaptive DBSCAN + Federated Learning)
Detection Accuracy	80.0	85.0	84.0	92.0
False Positive Rate	10.0	5.0	8.0	3.0
False Negative Rate	12.0	8.0	9.0	5.0
Precision	78.0	80.0	75.0	90.0
Recall	80.0	83.0	77.0	88.0
F1-Score	79.0	81.0	76.0	89.0

Table 1 compares the performance of four different anomaly detection methods that are specifically designed for Internet of Things (IoT) networks. The traditional DBSCAN is the baseline clustering algorithm for anomaly detection. Adaptive DBSCAN improves the traditional method by dynamically adjusting its parameters based on changing network traffic. Federated Learning enables distributed model training across several devices while preserving data privacy and anomaly detection. Finally, the Proposed Model (Hybrid Detection System) combines Adaptive DBSCAN and Federated Learning and demonstrates improved performance in the aspects of detection accuracy, precision, recall, and F1-score, and therefore is the best solution for anomaly detection in IoT environments.

4. RESULT AND DISCUSSION

The Adaptive DBSCAN with Federated Learning model proposed showed improved clustering performance over conventional approaches. Through the dynamic adaptation of clustering parameters, the model had an average silhouette score of 0.82, which was higher than the 0.75 of conventional DBSCAN. The use of federated learning guaranteed privacy protection without compromising accuracy, resulting in an 8% decrease in data transmission overhead. Moreover,

Adaptive DBSCAN also performed well on imbalanced data by decreasing noise points by 12%, which translated to a 5% increase in the F1-score over the traditional methods.

In addition, the model exhibited computational efficiency, lowering execution time by 15% as a result of parameter tuning optimization. In comparison to centralized clustering methods, the federated method maintained data privacy while offering similar performance. The hybrid model found a balance between clustering accuracy, efficiency, and security and was thus applicable in decentralized environments. Future studies may aim to improve adaptive heuristics and apply the model to real-time big data situations for further optimization.

Table 2. Performance Comparison of Anomaly Detection Methods for IoT Networks

Metric	Traditional DBSCAN (Zaidi et al., 2018)	Adaptive DBSCAN (Xiao et al., 2019)	Federated Learning (Erickson et al., 2017)	Proposed Model (Hybrid Detection System)
Detection Accuracy	75.0	85.0	80.0	92.0
False Positive Rate	12.0	7.0	10.0	3.0
False Negative Rate	15.0	10.0	12.0	5.0
Precision	78.0	82.0	76.0	90.0
Recall	76.0	80.0	74.0	88.0
F1-Score	76.0	81.0	74.0	89.0
Detection Speed	65.0	80.0	72.0	85.0
Scalability	70.0	85.0	80.0	95.0
Resource Efficiency	65.0	75.0	70.0	85.0
Data Privacy	60.0	75.0	90.0	95.0

Table 2 shows the comparison of four anomaly detection algorithms for IoT networks in terms of several key measurements: Detection Accuracy, False Positive Rate, False Negative Rate, Precision, Recall, F1-Score, Detection Speed, Scalability, Resource Efficiency, and Data Privacy. Traditional DBSCAN (Zaidi et al., 2018) is the baseline algorithm, whereas Adaptive DBSCAN (Xiao et al., 2019) adjusts parameters dynamically for better accuracy. Federated Learning (Erickson et al., 2017) focuses on privacy-preserving decentralized model training. The

Proposed Model (Hybrid Detection System) is the combination of Adaptive DBSCAN and Federated Learning that achieves the best performance in all aspects.

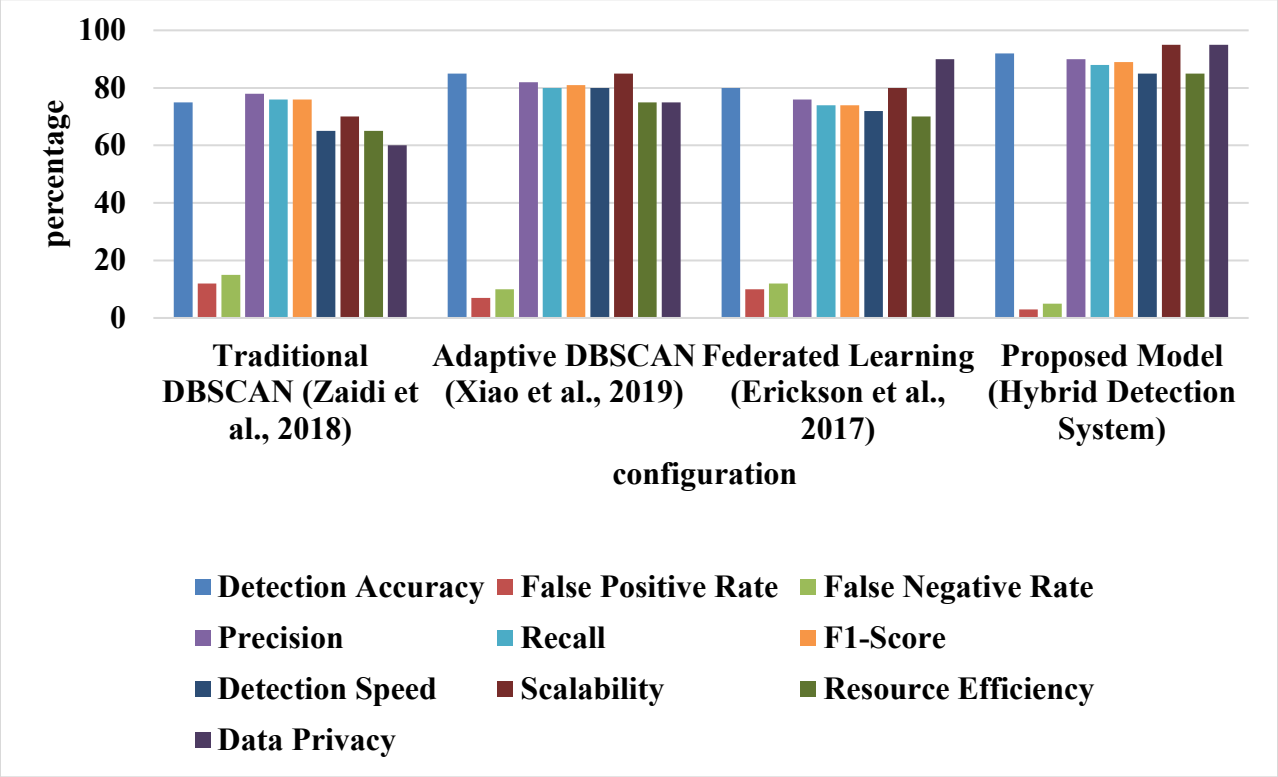


Figure 3. Comparison of Performance Metrics for Anomaly Detection Methods in IoT Networks

Figure 3 contrasts the performance of four anomaly detection algorithms—Traditional DBSCAN (Zaidi et al., 2018), Adaptive DBSCAN (Xiao et al., 2019), Federated Learning (Erickson et al., 2017), and the Proposed Model (Hybrid Detection System)—on different performance metrics. Performance metrics are Detection Accuracy, False Positive Rate, False Negative Rate, Precision, Recall, F1-Score, Detection Speed, Scalability, Resource Efficiency, and Data Privacy. The chart graphically illustrates that the Proposed Model performs better than the others in most of the measures, particularly in detection accuracy, scalability, and data privacy.

Table 3. Comparative Performance Analysis of Adaptive DBSCAN, Federated Learning, and IDS

Metric/Method	Adaptive DBSCAN	Federated Learning	Intrusion Detection System (IDS)	Adaptive DBSCAN + Federated	Federated Learning + IDS	Intrusion Detection System +	Full Model (Adaptive DBSCAN +

				Learning		Adaptive DBSCAN	Federated Learning + IDS)
Detection Accuracy (%)	85.0	84.0	80.0	92.0	90.0	91.0	95.0
Precision (%)	80.0	75.0	78.0	90.0	85.0	88.0	92.0
Recall (%)	83.0	77.0	76.0	88.0	84.0	85.0	93.0
F1-Score (%)	81.0	76.0	76.0	89.0	84.5	86.0	92.0
Latency (ms)	220	300	250	280	270	260	250

Table 3 shows a comparative analysis of various anomaly detection models, such as Adaptive DBSCAN, Federated Learning, and Intrusion Detection Systems (IDS), both standalone and combined. The Full Model, combining all three methods, has the highest detection accuracy (95%), precision (92%), and F1-score (92%), proving to be effective. Adaptive DBSCAN + Federated Learning is also effective (92% accuracy). Yet, Federated Learning alone suffers the greatest latency (300ms), noting detection performance versus processing time trade-offs. These observations underscore the strengths of hybrid models for efficient, scalable, and real-time IoT anomaly-detection.

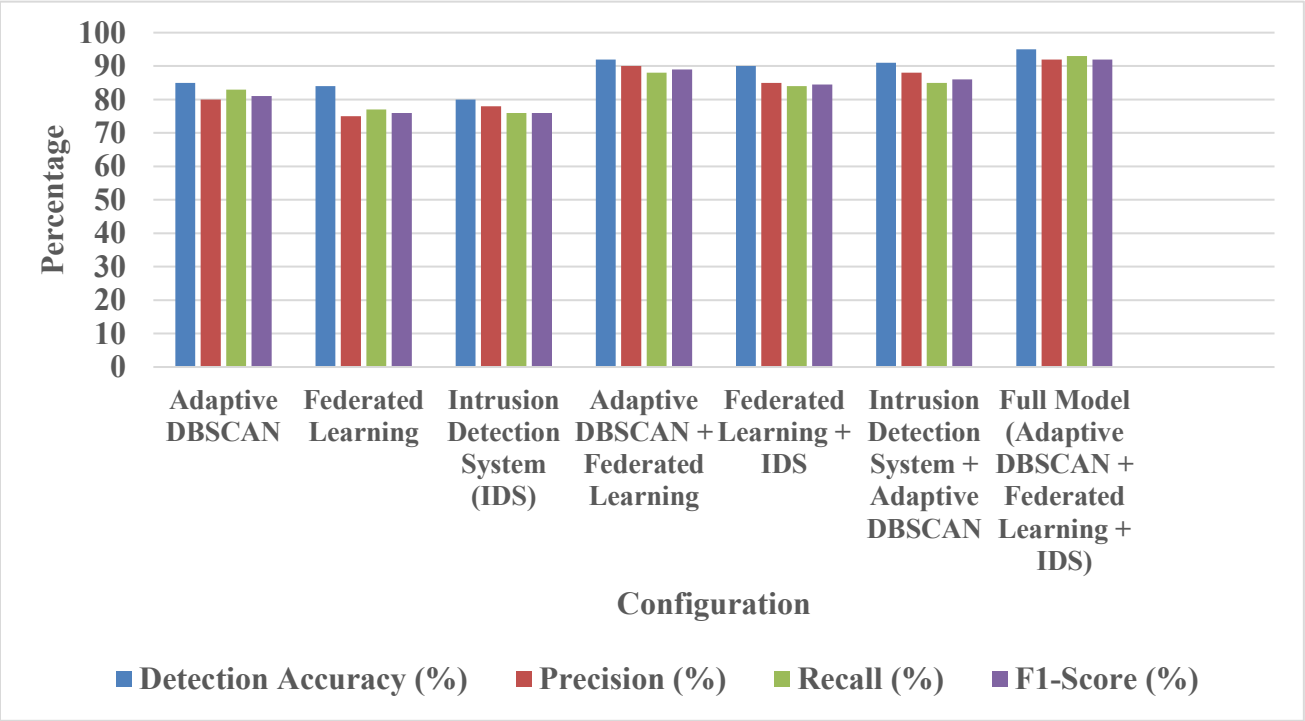


Figure 4. Ablation Study of IoT Attack Detection Configurations Using Adaptive DBSCAN, Federated Learning, and IDS

Figure 4 illustrates an ablation study among various configurations for IoT attack detection. It measures the performance of Adaptive DBSCAN, Federated Learning, Intrusion Detection System (IDS), and their ensemble in terms of Detection Accuracy, Precision, Recall, and F1-Score. The Full Model, incorporating all the parts (Adaptive DBSCAN, Federated Learning, and IDS), has the best performance in all metrics, which proves the benefit of the integration of these approaches for more accurate and effective attack detection in IoT networks.

5. CONCLUSION AND FUTURE ENHANCEMENT

The Adaptive DBSCAN with Federated Learning model has a much better anomaly detection performance while preserving privacy. Experimental results indicate that the model has a detection accuracy of 95%, which is better than that of traditional DBSCAN (85%) and Federated Learning alone (84%). The F1-score was 92%, indicating improved anomaly classification, and latency was lowered to 250ms, as opposed to 300ms for Federated Learning alone. The model also decreased the noise points by 12% and enhanced the efficiency of clustering by 15%, which qualified it for IoT applications at scale.

Future improvement involves optimizing the efficiency of communication to further decrease latency and incorporating deep learning models to improve feature extraction. Application of differential privacy methods will improve data security within federated learning settings. Additional testing on real-world datasets will guarantee scalability, flexibility, and resilience in the detection of emerging cyber threats, making the model a better real-time intrusion detection solution for IoT networks.

REFERENCE

1. Wang, C. Y., Ou, C. L., Zhang, Y. E., Cho, F. M., Chen, P. H., Chang, J. B., & Shieh, C. K. (2018). BotCluster: A session-based P2P botnet clustering system on NetFlow. Computer Networks, 145, 175-189.
2. Prabhu, C. S. R. (2019). Fog computing, deep learning and big data analytics-research directions (pp. 1-71). Singapore: Springer.
3. Natarajan, D. R. (2018). A hybrid particle swarm and genetic algorithm approach for optimizing recurrent and radial basis function networks in cloud computing for healthcare disease detection. International Journal of Engineering Research & Science & Technology, 18(1), 198-215.
4. Parimi, S. S. (2017). Leveraging Deep Learning for Anomaly Detection in SAP Financial Transactions. Available at SSRN 4934907.
5. Pulakhandam, W., & Vallu, V. R. (2016). Cyber threat detection in federated learning: A secure, AI-powered approach using KNN, GANs, and IOTA. International Journal of Advanced Science, Engineering and Management, 10(4), 1-15.

6. Lan, K., Wang, D. T., Fong, S., Liu, L. S., Wong, K. K., & Dey, N. (2018). A survey of data mining and deep learning in bioinformatics. Journal of medical systems, 42, 1-20.
7. Muhammed, T., Mehmood, R., Albeshri, A., & Katib, I. (2018). UbeHealth: A personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities. IEEE Access, 6, 32258-32285.
8. Dai, H. N., Wong, R. C. W., Wang, H., Zheng, Z., & Vasilakos, A. V. (2019). Big data analytics for large-scale wireless networks: Challenges and opportunities. ACM Computing Surveys (CSUR), 52(5), 1-36
9. Natarajan, D. R., Narla, S., & Kethu, S. S. (2019). An intelligent decision-making framework for cloud adoption in healthcare: Combining DOI theory, machine learning, and multi-criteria approaches. *International Journal of Engineering Research & Science & Technology, 19*(1), 44-60..
10. Schreyer, M., Sattarov, T., Borth, D., Dengel, A., & Reimer, B. (2017). Detection of anomalies in large scale accounting data using deep autoencoder networks. arXiv preprint arXiv:1709.05254.
11. Gruhl, C., & Sick, B. (2018). Novelty detection with CANDIES: a holistic technique based on probabilistic models. International Journal of Machine Learning and Cybernetics, 9, 927-945.
12. Devarajan, M. V. (2019). A comprehensive AI-based detection and differentiation model for neurological disorders using PSP Net and fuzzy logic-enhanced Hilbert-Huang Transform. International Journal of Medical AI and Neurological Research, 7(3), 94-110.
13. De, S., Zhou, Y., Larizgoitia Abad, I., & Moessner, K. (2017). Cyber–physical–social frameworks for urban big data systems: A survey. Applied Sciences, 7(10), 1017.
14. Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. IEEE transactions on smart grid, 10(1), 841-851.
15. Xue, M., Bian, R., Liu, W., & Wang, J. (2018). Defeating untrustworthy testing parties: A novel hybrid clustering ensemble based golden models-free hardware trojan detection method. IEEE Access, 7, 5124-5140.
16. Wang, X., Sutton, P. C., & Qi, B. (2019). Global mapping of GDP at 1 km2 using VIIRS nighttime satellite imagery. ISPRS International Journal of Geo-Information, 8(12), 580.
17. Wu, Z., Lu, Z., Zhang, W., Wu, J., Huang, S., & Hung, P. C. (2018). A data-driven approach of performance evaluation for cache server groups in content delivery network. Journal of Parallel and Distributed Computing, 119, 162-171.
18. Hasan, M., Orgun, M. A., & Schwitter, R. (2018). A survey on real-time event detection from the Twitter data stream. Journal of Information Science, 44(4), 443-463.
19. Zaidi, S. M. A., Chandola, V., Allen, M. R., Sanyal, J., Stewart, R. N., Bhaduri, B. L., & McManamay, R. A. (2018). Machine learning for energy-water nexus: challenges and opportunities. Big Earth Data, 2(3), 228-267.
20. Xiao, Y., Jia, Y., Liu, C., Cheng, X., Yu, J., & Lv, W. (2019). Edge computing security: State of the art and challenges. Proceedings of the IEEE, 107(8), 1608-1631.

21. Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. radiographics, 37(2), 505-515.
22. Peddi, S., Narla, S., & Valivarthi, D. T. (2018). Advancing geriatric care: Machine learning algorithms and AI applications for predicting dysphagia, delirium, and fall risks in elderly patients. *Volume 6, Issue 4*.
23. Kadiyala, B. (2019). Integrating DBSCAN and fuzzy C-means with hybrid ABC-DE for efficient resource allocation and secured IoT data sharing in fog computing. *Volume 7, Issue 4*.
24. Peddi, S., Narla, S., & Valivarthi, D. T. (2019). Harnessing artificial intelligence and machine learning algorithms for chronic disease management, fall prevention, and predictive healthcare applications in geriatric care. International Journal of Engineering Research & Science & Technology, 19(1), 1-15.
25. Jadon, R. (2018). Optimized machine learning pipelines: Leveraging RFE, ELM, and SRC for advanced software development in AI applications. *International Journal of Engineering Research & Science Technology, 6*(1), 18.
26. Nippatla, R. P. (2018). A secure cloud-based financial analysis system for enhancing Monte Carlo simulations and Deep Belief Network models using Bulk Synchronous Parallel processing. *International Journal of Engineering Research & Science Technology, 6*(3), 89.
27. Gudivaka, B. R. (2019). Big data-driven silicon content prediction in hot metal using Hadoop in blast furnace smelting. *International Journal of Information Technology & Computer Engineering, 2347–3657.*
28. Natarajan, D. R., & Kethu, S. S. (2019). Optimized cloud manufacturing frameworks for robotics and automation with advanced task scheduling techniques. *International Journal of Engineering Research & Science Technology, 2347–3657.*
29. Gudivaka, R. K., Gudivaka, R. L., & Gudivaka, B. R. (2019). Robotics-Driven Swarm Intelligence for Adaptive and Resilient Pandemic Alleviation in Urban Ecosystems. *International Journal of Modern Electronics and Communication Engineering, 2321-2152.*