Grenze International Journal of Engineering and Technology, June Issue

Intrusion Detection in Heterogeneous IoT Networks: A Future-based Comprehensive Study

Kamagari Shilppaa[1] and Suresh Kallam[2]
[1]Research Scholar, Department of Computer Science and Engineering, JAIN Deemed to be University, Bengaluru, India.
Email: k.shilppaa@jainuniversity.ac.in
[2]Department of Computer Science and Engineering, JAIN Deemed to be University, Bengaluru, India.
Email: suresh.k@jainuniversity.ac.in

***Abstract*— The Internet of Things (IoT) consists of physical objects that are associated by sensors, software, and connectivity to gather data and facilitate the transfer of data. The features that define IoT designs are heterogeneity, which presents a network with different devices, protocols, and networks with differing capacities, data formats. Intrusion Detection (ID) is crucial for cybersecurity, as it provides an effective means of real-time monitoring systems and networks to respond to malicious activities, anomalies, and violations of policy from a security perspective in heterogeneous IoT. This survey examines various ID techniques, including Deep Learning (DL), Machine Learning (ML), Rule-based, and Statistical, Blockchain-enabled Intrusion Detection System (IDS) techniques for detecting intrusions. The study highlights new trends, including federated IDS, explainable Artificial Intelligence (AI), and lightweight edge security. This also discusses significant challenges that need to be tackled in order to create effective, scalable, and adaptable ID solutions for future IoT environments.**

***Index Terms*— Internet of Thing, Deep Learning, Machine Learning, Heterogeneity, Heterogeneous IoT, Intrusion Detection System.**

I. INTRODUCTION

The Internet of Things (IoT) represents an interrelated computing sensors, devices, and media which can exchange data and work together to connect and interact over the Internet. IoT connects different smart technologies, such as smart cities, smart homes, and smart energy plans, make them broadly use [1]. Nowadays, IoT devices have a wide range of uses. As IoT networks become more widespread, the usage of different heterogeneous IoT networks such as cellular, Wi-Fi, Bluetooth low energy, and others, will be rising on a daily basis [2]. Whereas nearby several industries that can help from the usage of IoT devices, these devices also be a major source of cyberattacks with recent incidents demonstrating significant financial losses as an outcome of exploitable susceptibilities in IoT devices [3]. As technology continues to grow, the number of security vulnerabilities continues to increase. Consequently, in fighting these new threats, Intrusion Detection System (IDS) will be necessary. IDS has been a necessity in securing IoT devices and networks from a diversity of threats [4].

In view of the discussion on existing surveys, this study focuses on the following

- Comparative assessment of ML, DL, statistical, rule-based and blockchain IDS approaches under IoT heterogeneity constraints.

Grenze ID: 01.GIJET.12.2.277_21

- A new conceptual IDS framework integrating lightweight edge detection, adaptive cloud learning, and blockchain-based trust sharing.
- Cross-layer analysis linking IoT device diversity, protocol variability, and resource imbalance to detection performance.
- This survey clearly defined research directions for concept-drift handling, energy-aware model development, and federated privacy-preserving IDS.

The following structure will discuss the organization of this paper: In Section 2, provide a characterization of Heterogeneous IoT Networks. Section 3 outline the Fundamentals of IDS Architectures. Section 4 presents the Taxonomy of Intrusion Detection (ID) in Heterogeneous IoT. Section 5 provide Datasets and Evaluation Metrics. Section 6 examine the Challenges in ID for Heterogeneous IoT. Section 7 explain the Comparative Analysis of Existing Approaches. Section 8 represent the Emerging Trends and Future Directions. Section 9 indicate the Open Research Challenges. Lastly, conclude this study in Section 10.

II. Overview of Heterogeneous IoT Networks

This section covers the overview of Heterogeneous IoT networks, including characteristics, architecture, challenges, and security issues.

A. Characteristics of Heterogeneous IoT

The characteristics of Heterogeneous IoT such as device diversity, protocol diversity, data format variability, and cross-layer communication, are elaborated in Table I.

Table I. Characteristics of Heterogeneous IoT

Characteristics	Description
Device Diversity	The IoT systems include devices like sensors, smartphones and the devices that vary in hardware, operating system, vendors and manufacturers.
Protocol Diversity	The device capabilities and use cases are used in different communication protocols.
Data format Variability	The network handles information from numerous sources, leading to considerable diversity in data types, formats, and structures.
Cross-Layer Communication	Enhancing performance, increasing efficiency, and addressing complex network needs, particularly in IoT and wireless systems, wherever resources are limited and conditions are constantly changing.

B. Security Concerns in Heterogeneous IoT

The nature of the heterogeneous IoT necessitates reliable communication that requires advanced and scalable security measures. Backbone networks must gather more data due to the huge number of heterogeneous IoT devices. Because the constraints of heterogeneous network components vary greatly and is essential to have an efficient heterogeneous topology structure to arrange all of the components for communication and optimize their effectiveness. There is heterogeneity and extra connectivity issues since dissimilar applications use variety network designs [5].

C. Attacks Mapped to IDS Detection Techniques

DL-based IDS using CNN–LSTM can detect abnormal outbound request bursts that precede botnet propagation. Rule-based IDS detects unauthorized login patterns, but fails in credential-stuffing zero-day scenarios where anomaly-based ML performs better. Blockchain-enabled IDS can secure telemetry sharing across hospital IoT infrastructure, preventing lateral intrusion propagation. Smart home Botnet attacks detect sudden traffic deviations, but ML-based feature learning improves classification when multi-device traffic is involved.

III. Taxonomy of ID in Heterogeneous IoT

The IDS plays a role in network security by identifying malicious or unauthorized actions and enabling prompt responses. An IDS can be characterized according to elements like data origin, system architecture, and detection methods.

A. Based on Detection Approach

In this section, discuss the various detection approaches in IDS. The ML-based, DL-based, Statistical-based and Rule-based, and Blockchain-enabled IDS approaches are used to detect the intrusion. Table 2 lists the comparison of IDS detection approaches to analyse the benefits and limitations of the survey.

- ML-based: ML-based IDS approaches focus on learning discriminative patterns from network traffic to identify normal and malicious behavior. ML approaches generalize attack patterns using a lightweight algorithm such as decision trees, random forest, Support Vector Machine (SVM), and naïve bayes. However, ML algorithms are highly dependent on the quality of manually engineered features. This ML-based approaches perform in complex networks but degrade under high protocol diversity due to manual feature dependency and concept drift sensitivity. They struggle to adapt to concept drift, where attack behaviors evolve over time, resulting in decreased detection performance.
- DL-based: DL-based IDS approaches utilize multi-layer neural networks such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and autoencoders to automatically learn high-level representations. The DL-based approaches needed high resources, unsuitable for low-power IoT edge nodes. It effectively captures temporal and spatial traffic patterns and detects previously unseen attack signatures. However, the primary limitation of DL-based IDS is its computational demand, training and inference require significant memory, processing power, and time.
- Statistical-based: The statistically based IDS utilizes two separate profiles: one profile is for statistical training while the other one is for tracking ongoing network activity. If the anomaly score increased beyond a predetermined threshold, the IDS will generate an alert indicating a possible intrusion. These methods rely on behavioral patterns to classify observations as outliers in the dataset.
- Rule-Based: This method is used with both the anomaly method and signature method. Malicious attacks are detected in signature detection by examining packets based on the predefined rules of the system. Anomaly detection, examines the behavior of the system and notes the difference experience between abnormal behavior and normal behavior based on predetermined rule set, such as the order in which programmers call the system. [8]. This well suitable for known-attack identification but fail in zero-day scenarios and require constant rule updates.
- Blockchain-enabled IDS: IDSs have also leveraged the power of blockchain technology to identify malicious attacks. Blockchain holds the difference to the difficulties with the anomaly and signature methods of detection, as the underpinning foundation of the blockchain-based IDS models. This introduces network overhead, making them more suited for fog/cloud layers rather than constrained edge devices

TABLE II: ANALYSIS OF IDS DETECTION APPROACHES IN HETEROGENEOUS IOT

IDS Approach	Benefits	Limitations	Best Use Scenarios
ML-based IDS	Good interpretability, fast training, and works with engineered features	Sensitive to concept drift, requires manual feature engineering, and prone to bias in imbalanced datasets	Lightweight detection, structured data, and moderate-scale networks
DL-based IDS	Learns complex non-linear patterns, no manual feature engineering, and high detection accuracy	High computation cost, low interpretability, and not fully edge-friendly without optimization	Advanced attack detection, large-scale IoT traffic, and cloud-driven security
Statistical-based IDS	Lightweight, simple implementation, and fast decision output	Inflexible to evolving threats, and high false alarms with dynamic traffic	Baseline anomaly detection, fixed-pattern environments, and resource-constrained nodes
Rule-based IDS	Very low overhead, precise for known attacks, and easy rule deployment	Cannot detect zero-day threats and requires continuous rule updates	Known attack signatures, small networks, and pre-authorized device environments
Blockchain-enabled IDS	Tamper-proof audit trails, trust sharing between nodes, and no single point of failure	Higher latency, storage overhead, and consensus energy cost	Threat intelligence sharing, multi-domain IoT, and collaborative detection networks

B. Based on Deployment Environment

This section covers the Edge/Fog-level detection, Cloud-level analysis, and Collaborative or Federated IDS.

- Edge/Fog-level detection: The Edge/Fog layer is the most important because it contains the edge/fog nodes, which executes near-line processing and analysis to promote real-time responsiveness. The networking hardware is able to guarantee consistent and active data transmission and communication between devices and nodes.
- Cloud-level analysis: Cloud layer is in charge of monitoring the overall global system. Because the cloud layer involves computing larger sizes of processing (compared to the edge layer), ML algorithms can run more difficult in finding patterns in how the system is operated. Cloud layer can also improve inference of the parameters utilized in the edge layer models. Therefore, that cloud layer can recommend the edge layer to update its ML advance models, to improve the performance of the model.

- Collaborative or Federated IDS: Collaborative IDS (CIDS) for accurate and efficient ID in a distributed system and Federated IDS both provide valid and distinct ways to provide ID, and vary in the ways that they address data and privacy models. CIDS, shares aggregated threat data directly, while Federated IDS uses ML to share model updates without sharing original data the actual sensitive data.

IV. Datasets and Evaluation Metrics

In this study, we focused on different evaluation metrics employed in the papers.

A. Benchmark Datasets

Standardized high-quality datasets known as benchmark datasets are utilized to fairly evaluate and compare the models that perform IDS.

NSL-KDD (2009) dataset: It was developed by Ebrahim Bagheri, Mahbod Tavallaee, Ali A. Ghorbani, Wei Lu. It has 41 features. There are no recent attacks and a reasonable size of class-unbalanced network traffic with redundant and duplicate samples eliminated. The dataset lacks modern IoT signatures and limiting real-world scenarios.

CICIDS2017 dataset: It was developed by Iman Sharafaldin, Arash Habibi Lashkari, and Ali A. Ghorbani, and it has 80 features. This dataset is flow-based and includes meta-data about IP addresses and attacks [9]. The dataset includes rich traffic but has imbalanced attack distribution.

BoT–IoT Dataset: Specifically created for assessing IoT botnet detection the BoT–IoT is a diverse multi-class labeled dataset that is produced in a real-world testbed environment. Furthermore, a Testing and Training network record file with 45 unique features is involved in the dataset. However, the dataset includes only limited types of attacks.

ToN_IoT Dataset: The TON IoT dataset, displayed by Saedi et al., is made from an Industry 4.0/IIoT testbed comprising of three layers, counting edge, fog, and cloud, to simulate the real-world IoT environment. In expansion, the dataset too contains a Preparing and Testing organize record that incorporates 45 particular highlights. However, the dataset faced limitations from label noise and feature sparsity which complicates training stability.

IoTID20 Dataset: The IoTID20 dataset will provide a better way to detect security threats in IoT networks. The IoTID20 data can be available in CSV file. It includes 12 features. The dataset includes limited features, which affects the IDS performance.

V. Challenges in ID for Heterogeneous IoT

The section, discuss the Challenges in ID for Heterogeneous IoT.

A. Scalability and Resource Constraints

These schemes have been applied so far without any performance degradation but handle a variation of devices and escalating data loads. The main challenges include scalability and resource constraints: limited hardware usage, energy efficiency, vulnerability to resource exhaustion, and delayed [10].

B. Real-Time Detection and Latency

The detection layers required for real-time ID and several latency reduction techniques represent one complicated procedure, the heterogeneity of the network including a number of devices and protocols makes it more than not easy to achieve the low-latency response able to prevent the evolving cyber-attacks [11].

C. Data Imbalance and Concept Drift :

The concept drift and data imbalance happen at the similar time, which indicates to ML models becoming inaccurate over time in heterogeneous IoT ID. The concept drift is the term used to describe how attack patterns or typical behaviors gradually change, making prior knowledge outdated, whereas data imbalance shows that attacks are rare compared to regular traffic. [12, 13].

VI. Emerging Trends and Future Directions

The blockchain and its many features have improved security performance over peer-to-peer networks without requiring a third party. A thorough analysis of recently developed blockchain-based IDS is also provided.

A. Federated and Distributed IDS Models

Real-time applications can nowadays more simply usage FL-based IDS thanks to the communication-efficient FL systems like Dynamic Weighted Aggregation Federated Learning (DAFL) which, lower training costs and network latency. To solve issues of scalability, collaboration, trust, and security, emerging trends include the use of blockchain and peer-to-peer architectures. Blockchain offers distributed consensus, tamper-proof data storage, and improved security for applications such as the IoT [14, 15].

B. AI and Explainable IDS Models

XAI solves this problem by making the rationale behind an AI model's output transparent, hence allowing for further development of more transparent, reliable, and efficient IDS models. In future work, an app will be proposed that will implement an idea of explainable AI on different complex datasets such as ToN_IoT for real-time data investigation and evaluation of performance of predictions [16].

C. Lightweight IDS for Edge Devices

Hybrid solutions for lightweight IDS can be considered, taking in resource-efficient methods for pruning and quantization to achieve performance with the help of DL models such as CNNs and BiLSTMs to ensure accuracy. An effective way in which they work together is that they provide a hybrid model that perfectly balances low latency with high accuracy at complexity with the datasets [17, 18].

VII. Open Research Challenges

The section includes challenges involved in IDS.
The efficiency of AIDS analysis in identifying cyberthreats is carefully evaluated in comparison to innovative approaches. During a six-month test stage quantitative data is collected to examine constraints by way of prediction rate FP, FN and occurrence response time. As per proved by the important enhancement in finding and the discount in FP and FN the results display. This ensures AIDS is superior due to its capability to adapt and respond to new threats [19-21].

A. Concept irift

Attack patterns change continuously, where static ML/DL models decrease up to 30-45% accuracy within long-term deployments.

B. Class imbalance

The CICIDS2017 and IoTID20 datasets contain high legitimate-to-attack ratios, causing ML/DL models to favor normal traffic. There is need for cost-sensitive learning, SMOTE-based resampling, and hybrid feature balancing approaches.

VIII. Conclusion

The survey examines the IDS can support with regulatory compliance, enhance network visibility, and detect threats early. This study mainly aims to detect intrusions, protect data, and ensure its security and highlights analytical comparison to improve the originality of the survey. The system would be beneficial to enable consistent comparison of performance results across different research efforts. This is to detect the accurate behaviors of IoT networks and validate the results of research endeavors. Even though detection tools for accurately identifying known attacks take finished important development, there are still issues like the huge rate of FP, updating the model, and detecting zero-day attacks. Future studies might concentrate on addressing these issues. Improvements in the comparison of ML models are therefore regarded as a promising field of study. In order for organizations to take the proper control measures to mitigate risk proactively, this is also necessary to bridge the gap between incident management and cybersecurity systems.

References

[1] P. Venkatesh, and Mr M. Manikandan, "The Role of The Internet of Things in The Telecom Sector,".

[2] K. Shilppaa and S. Kallam, "A Quantum-Resistant Cryptography for Secure IoT Communication Using Novel QRC_BIKAES Framework," 2025 IEEE 5th International Conference on VLSI Systems, Architecture, Technology and Applications (VLSI SATA), Bangalore, India, 2025.

[3] M. Fatima, O. Rehman, I. M. Rahman, A. Ajmal, and S. J. Park, “Towards ensemble feature selection for lightweight intrusion detection in resource-constrained iot devices,” Future Internet, vol. 16, pp. 368, Oct 2024.

[4] S. B. Sharma, and A. K. Bairwa, “Leveraging AI for Intrusion Detection in IoT Ecosystems: A Comprehensive Study,” IEEE Access, Mar 2025.
[5] S. S. Mahadik, P. M. Pawar, and R. Muthalagu,” Heterogeneous IoT (HetIoT) security: techniques, challenges and open issues,” Multimedia Tools and Applications, vol. 83, pp. 35371-35412, Apr. 2024.
[6] V. Shanmugam, R. Razavi-Far, and E. Hallaji, “Addressing class imbalance in intrusion detection: a comprehensive evaluation of machine learning approaches,” Electronics, vol. 14, pp. 69, Dec. 2024.
[7] S. S. Alsaleh, M. E. B. Menai, and S. Al-Ahmadi, “Federated Learning-Based Model to Lightweight IDSs for Heterogeneous IoT Networks: State-of-the-Art, Challenges, and Future Directions,” IEEE Access, vol.12, pp. 134256-134272, sep. 2024.
[8] T. Al-Shurbaji, M. Anbar, S. Manickam, I. H. Hasbullah, N. ALfriehate, B. A. Alabsi, A. R. Alzighaibi, and H. Hashim, “Deep learning-based intrusion detection system for detecting IoT botnet attacks: a review,” IEEE Access, Jan. 2025.
[9] N. Gupta, V. Jindal, and P. Bedi, “A survey on intrusion detection and prevention systems,” SN Computer Science, vol. 4, vol. 439, June 2023.
[10] M. Kamdjou, D. Baudry, V. Havard, and S. Ouchani, “Resource-constrained extended reality operated with digital twin in industrial internet of things,” IEEE Open Journal of the Communications Society, vol. 5, pp. 928-950, Jan. 2024.
[11] M. Muñoz, M. Torres, J. D. Gil, and J. L. Guzmán, “An Internet of Things platform for heterogeneous data integration: Methodology and application examples,” Journal of Network and Computer Applications, pp. 104197, Apr. 2025.
[12] K. M. Junaid, D. Paulraj, and T. Sethukarasi, “A comprehensive ensemble classification techniques detecting and managing concept drift in dynamic imbalanced data streams,” Wireless Networks, vol. 31, pp. 19-30, Jan. 2025.
[13] M. K. Dave, and B. K. Mittapally, “Data integration and interoperability in IoT: challenges, strategies and future direction,” Int. J. Comput. Eng. Technol. (IJCET), vol. 15, pp. 45-60, Jan. 2024.
[14] B. Buyuktanir, S. Altinkaya, G. Karatas Baydogmus, and K. Yildiz, “Federated learning in intrusion detection: advancements, applications, and future directions,” Cluster Computing, vol. 28, pp. 1-25, Sep. 2025.
[15] W. Li, C. Stidsen, and T. Adam, “A blockchain-assisted security management framework for collaborative intrusion detection in smart cities,” Computers and Electrical Engineering, vol. 111, pp. 108884, Oct. 2023.
[16] I. Samed, and S. Sagiroglu, “Explainable artificial intelligence models in intrusion detection systems,” Engineering Applications of Artificial Intelligence, vol. 144, pp. 110145, Mar. 2025.
[17] N. Gaddam, “AI-based intrusion detection system for edge computing security,” machine learning (iscsitr-ijsraiml), vol. 4, pp. 1-15, Jan. 2023.
[18] Y. Baseri, V. Chouhan, and A. Hafid, “Navigating quantum security risks in networked environments: A comprehensive study of quantum-safe network protocols,” Computers & Security, vol. 142, pp. 103883, Jul. 2024.
[19] B. R. Yadav, “Machine learning algorithms: optimizing efficiency in AI applications,” International Journal of Engineering and Management Research, vol. 14, pp. 49-57, Jul. 2024.
[20] S. Jamshidi, A. Nikanjam, K. W. Nafi, F. Khomh, and R. Rasta, “Application of deep reinforcement learning for intrusion detection in Internet of Things: A systematic review,” Internet of Things, pp. 101531, Feb. 2025.
[21] M. Barbu, A. V. Vevera, and D. C. Barbu, “Standardization and interoperability—Key elements of digital transformation,” In Digital transformation: Technology, tools, and studies (pp. 87-94). Cham: Springer Nature Switzerland, Apr. 2024.