EAI Endorsed Transactions
on Internet of Things

Research Article EAI.EU

A Q-Learning and Blockchain Framework for Secure Dynamic Bandwidth Allocation in Heterogeneous IoT

Ahmed Saeed Obied [1 2*], Ahmad Shahidan Abdullah[1], Hind Mowafaq Taha[3], Sanaa Abd Tarish[4]

[1] Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia
[2] Public Law Department, College of Law, Al-Nahrain University, Baghdad, Iraq
[3] Electronic and Communications Engineering Department, College of Engineering, Al-Nahrain University, Baghdad,Iraq
[4] Public Law Department, College of Law, Al-Iraqia University, Baghdad, Iraq

Abstract

The rapid expansion of the Internet of Things (IoT) has intensified the challenge of achieving dynamic bandwidth allocation while maintaining security across heterogeneous devices and communication protocols. Conventional static allocation schemes lack adaptability, while existing learning-based or blockchain-based approaches typically optimize performance or trust in isolation. To address this gap, this paper proposes a hybrid framework that integrates Q-learning–based adaptive bandwidth allocation with a lightweight, permissioned blockchain-based trust mechanism. The framework is evaluated through MATLAB-based simulations involving 100 heterogeneous IoT devices under dynamic traffic conditions and adversarial behavior. Performance is compared against multiple baselines, including static allocation, learning-only and blockchain-only schemes, classical scheduling algorithms (WFQ and DRR), and a deep reinforcement learning approach (DQN). The results reveal clear trade-offs among bandwidth utilization, fairness, energy consumption, and security. Static and classical schedulers provide predictable fairness but remain vulnerable to malicious activity. Learning-only and deep reinforcement learning approaches improve adaptability but lack intrinsic trust awareness, while blockchain-only enforcement enhances security at the expense of responsiveness. By coupling adaptive decision-making with trust validation, the proposed hybrid framework achieves a balanced operating point, offering stable bandwidth utilization, improved energy efficiency, and robust attack resilience under noisy and uncertain conditions. These findings highlight the importance of aligning learning mechanisms with trust-aware constraints for secure and scalable bandwidth management in heterogeneous IoT networks.

Keywords: Heterogeneous IoT, Dynamic Bandwidth Allocation, Machine Learning, Blockchain, Network Security.

Received on 04 October 2025, accepted on 14 December 2025, published on 03 February 2026

doi: 10.4108/eetiot.10464

*Corresponding author. Email: ahmad_happy2005@yahoo.com

1. Introduction

The Internet of Things (IoT) has rapidly evolved from a conceptual vision into a mainstream technological reality, with billions of devices currently connected and continuously exchanging data, often with little or no human intervention [1]. Recent reports indicate that the number of IoT devices surpassed 15 billion in 2023 and is projected to reach over 29 billion by 2030, making IoT one of the fastest-growing infrastructures in modern digital ecosystems. This exponential growth is transforming industrial sectors, healthcare, transportation, and smart cities, but it also raises significant challenges in resource management and network security. As IoT networks expand, heterogeneity emerges as a fundamental challenge. Devices employ diverse communication protocols (e.g., ZigBee, Wi-Fi, LoRa, NB-IoT), exhibit different latency requirements, and operate under varying levels of reliability and trust [2]. Managing such diversity makes bandwidth allocation increasingly complex, especially when considering both efficiency and security.

Traditional static bandwidth allocation methods fail to cope with dynamic traffic variations, leading to resource underutilization, congestion, and increased vulnerability to cyberattacks such as denial-of-service (DoS) and bandwidth hijacking [3, 4]. Likewise, conventional security frameworks designed for homogeneous networks are inadequate in heterogeneous and large-scale IoT environments.

To address these limitations, researchers have increasingly turned to Artificial Intelligence (AI). In particular, Machine Learning (ML) and Deep Reinforcement Learning (DRL) have been explored to improve adaptability, optimize throughput, and minimize latency in bandwidth management [5]. For example, edge-based AI systems have demonstrated real-time decision-making and efficient resource utilization, achieving latency reductions of up to 46% in certain scenarios [6]. However, most AI-driven solutions prioritize performance at the expense of holistic security.

Conversely, blockchain has emerged as a decentralized trust mechanism, offering transparency, immutability, and resilience against malicious activity [7]. While blockchain enhances device-to-device trust and data integrity, its energy and computational overheads make it less feasible for resource-constrained IoT deployments.

Although AI and blockchain are individually powerful, their combined integration remains underexplored. Existing studies often focus on performance-driven AI methods or security-driven blockchain solutions in isolation. A limited number of attempts have combined both, but these are typically constrained to small-scale simulations and fail to address scalability, adaptability, and practical deployment challenges.

This study addresses these gaps by proposing a unified framework that integrates Q-learning–based adaptive bandwidth allocation with blockchain-inspired trust management. The key contributions of this paper are as follows:

- Propose a unified trust-aware bandwidth allocation framework that tightly integrates Q-learning–based adaptive decision-making with a lightweight, permissioned blockchain-inspired trust validation mechanism for heterogeneous IoT networks.
- Design a learning process in which trust validation is embedded directly within the Q-learning loop, allowing bandwidth allocation decisions to adapt dynamically while being progressively constrained by observed device behavior under adversarial and uncertain conditions.
- Conduct a comprehensive evaluation using MATLAB-based simulations, comparing the proposed framework against static allocation, AI-only, blockchain-only, classical scheduling algorithms, and deep reinforcement learning baselines, thereby demonstrating the practical trade-offs between adaptability, energy efficiency, fairness, and security.

To contextualize the research gap, Table 1 presents a comparative analysis of representative studies, emphasizing the shortcomings of static methods, the partial focus of AI-only techniques, and the high overheads of blockchain-only solutions. These observations highlight the necessity for the integrated hybrid framework proposed in this work.

Table 1. Comparative analysis of existing approaches.

Reference	**Bandwidth Allocation**	**Security & Privacy**	**IoT Attacks**	**Threat Models**	**Blockchain in IoT**	**Q-Learning in IoT**	**Key Areas Covered**
Zhou et al., 2020 [6]	✓	✗	✗	✗	✗	✗	RL-based bandwidth allocation in NG-EPON
Arshad et al., 2023 [8]	✗	✓	✓	✗	✓	✗	Blockchain-based decentralized trust in IoT
Obaidat et al., 2024 [9]	✗	✓	✓	✓	✓	✗	Comprehensive IoT + Blockchain survey
Haque et al., 2024 [10]	✗	✓	✗	✗	✓	✗	Lightweight blockchain consensus for IoT
Hatem et al., 2019 [12]	✓	✗	✗	✗	✗	✗	DL-based dynamic bandwidth allocation for optical access
Wong & Ruan, 2023 [13]	✓	✗	✗	✗	✗	✗	Self-adaptive bandwidth allocation for 6G fronthaul
Liem et al., 2023 [15]	✓	✗	✗	✗	✗	✗	LSTM-based dynamic bandwidth allocation
Proposed work, 2025	✓	✓	✓	✓	✓	✓	Unified framework: Q-learning + Blockchain for secure dynamic bandwidth allocation

Significance. This integrated approach not only improves resource utilization but also strengthens IoT resilience against malicious activities, providing a scalable foundation for future 6G and beyond networks. To the best of our knowledge, this study is among the first attempts to unify reinforcement learning with blockchain-based trust management into a practical bandwidth allocation framework validated through simulation. Unlike prior studies that typically address either bandwidth optimization e.g. [6, 12, 15], or blockchain-based trust e.g. [8- 10], our work integrates both aspects into one unified model.

As summarized in Table 1, our work addresses both dynamic bandwidth allocation and secure trust management in an integrated manner. The following section reviews related literature in more detail.

The remainder of this paper is structured as follows. Section 2 reviews related literature on bandwidth allocation, IoT security, and heterogeneous network management, highlighting existing gaps. Section 3 presents the background of the proposed framework, including IoT heterogeneity, Q-learning fundamentals, blockchain principles, trust models, and the mathematical formulation. Section 4 describes the design of the integrated architecture, its three-layer structure, workflow, and the simulation setup and parameters used for evaluation. Section 5 discusses the experimental results and comparative analysis across static, AI-only, blockchain-only, and hybrid scenarios. Finally, Section 6 concludes the study and suggests directions for future research.

2. Related Work

This research focuses on three main themes: (1) Bandwidth allocation methods, (2) Security considerations in bandwidth management, and (3) Dealing with heterogeneity in IoT systems. Each area has played a key role in shaping the proposed framework.

2.1 Bandwidth Allocation: Static and Dynamic

2.1.1 Static vs Dynamic Approaches

Most bandwidth in IoT networks remains subject to static policies. These methods are easy to apply, but in practice they cannot cope with the continuously changing IoT traffic patterns. Loads, latency needs, and device diversity all change on an ongoing basis [11].

As a result, static allocation wastes bandwidth, causes congestion, and makes the system susceptible to attack. Dynamic approaches, by contrast, have come into view recently as a better choice. Intelligent algorithms are used to adjust resources in real time and make the network more efficient [6].

2.1.2 AI and Machine Learning Techniques

In recent years, people began to use artificial intelligence (AI) and machine learning (ML) more and more to deal with bandwidth. One case is the PABA framework [11]. The idea is simple: edge devices train together and then share one model. After that, the bandwidth is divided by looking at how strong each device is. Tests said delay went down by almost half (46%) and accuracy moved up a bit (about 4%). This means it can work even when the system does not have enough resources.

In optical networks, deep learning (DL) was tried with NG-EPON. Here, the model tried to guess how much bandwidth might be needed next, and this helped to cut down on control cost [12]. Another method, called JAAPD-D [14], mixed AI ideas with Lyapunov theory. It managed delay, computing, and communication all at once. From the tests, the system became faster and more efficient.

Some works also used LSTM. They showed that video traffic runs smoother, with less delay and fewer jumps. In the end, the quality for users got better [15]. All of these studies tell us one thing: ML can adapt the bandwidth much better than the old fixed rules.

2.1.3 Reinforcement Learning-based Approaches

Reinforcement Learning (RL) is becoming more prominent in the allocation of bandwidth for IoT devices as it can adaptively learn from different network scenarios and traffic trends. RL-based models in 6G networks adapt to varying traffic needs, offering low latency for sensitive applications [13].

In Software-Defined Networking (SDN) environments, Deep Q-Learning (DQL) has been applied to intent-based routing. This enables self-driven bandwidth allocation that increases throughput and supports efficient path switching [16]. In contrast, contextual multi-armed bandit approaches have been suggested for millimeter-wave (mmWave) 5G systems. These methods allow adaptive resource allocation that changes with channel conditions while balancing exploration and exploitation [20].

2.1.4 Alternative and Hardware-Enhanced Methods

Besides AI-based solutions, other rule-based methods have also been explored. A good example is the Adaptive Hungarian Algorithm (AHA), which allocates bandwidth in network slicing scenarios with little need for training. In practice, it achieves throughput close to ML-based methods, but with a fraction of the computational cost [17].

At the hardware level, innovative designs such as dual-memory pad architectures have been proposed to cut energy use and improve memory bandwidth in edge devices [18]. In Mobile Edge Computing (MEC), we humbly acknowledge the study of bandwidth-aware scheduling using Pareto optimization and hybrid CPU–GPU scheduling for better scalability and energy efficiency [19]. We also recognize the exploration of backend allocation for on-device AI inference to enhance performance with limited resources [21].

2.1.5 Summary

Previous studies have shown a movement from fixed bandwidth allocation to dynamic approaches with AI. While ML and DL approaches enable predictive bandwidth management, RL techniques offer real-time adaptability in highly variable environments. Alternative rule-based solutions and hardware-enhanced methods further complement these strategies by addressing energy efficiency and computational scalability.

Table 2 presents various methods for bandwidth allocation that utilize AI, thoughtfully examining the pros and cons of each approach, alongside the intended applications.

Table 2. Summary of AI-Driven Dynamic Bandwidth Allocation Techniques in IoT Networks.

Paper	Year	Domain	AI/ML Technique	Primary Metric	Performance Gain
Wen, D [11]	2020	Edge Learning	Optimization	Latency, Accuracy	↓ 46% Latency, ↑ 4% Accuracy
Hatem, J [12]	2019	NG-EPON	Deep Learning	Control Overhead	↓ Overhead, ↑ Bandwidth Utilization
Wong, E [13]	2023	6G MFH	Reinforcement/Transfer Learning	Latency	Rapid adaptation
Hu, Y et al [14]	2024	6G AIaaS	DNN, Lyapunov Opt.	Delay, Throughput	↓ Delay, ↑ Throughput
Liem, A et al. [15]	2023	NG-EPON	LSTM	Delay, Jitter, BW Util.	↓ Delay, ↓ Jitter, ↑ BW Util.
Żotkie et al. [16]	2021	SDN, IoT	Deep-Q-Learning	Throughput, Utilization	↑ Throughput, ↑ Utilization
Chen, Y [17]	2023	Network Slicing	AHA (Rule-based)	System Throughput	93-97% Max. Throughput
Chen, w [18]	2024	Edge AI Hardware	Custom Arch. (CIM, DLMP)	Area/Power Eff.	207.4 GOPS/mm^2, 3.53 TOPS/W
Lin, Z [19]	2020	MEC AI Service	CNN, DNN, Search	Comp. Time, Energy	Near-optimal performance
Qureshi, M [20]	2020	5G mmWave	MAB (Unimodal)	Regret	Logarithmic regret
Iyer, V et al. [21]	2023	On-Device AI	Feedback, Pareto	Throughput	↑ 25-100% Throughput

2.2 Security Issues in Bandwidth Allocation

The Internet of Things (IoT)'s remarkable expansion has overwhelmed bandwidth distribution methods with significant weaknesses, making networks susceptible to a range of cyber threats such as DoS and bandwidth appropriation. The offensives compromise the soundness, availability, and reliability of IoT frameworks, establishing secure and efficient bandwidth management as an essential need. Established models of security, which are usually tailored for steady or uniform systems, prove inadequate in the realm of fast-evolving and varied systems like IoT architectures [22].

2.2.1 Bandwidth Allocation Cybersecurity Threats

A diverse array of security assaults challenges IoT networks, directly affecting bandwidth regulation. DoS and distributed DoS (DDoS) threats can overwhelm communication connections and render the service inaccessible. Bandwidth hijacking allows malware to exploit excessive resource usage by misusing legitimate device resources, and spoofing and Sybil attacks aid attackers in creating fake identities and grabbing bandwidth allocations. Also, eavesdropping and data interception violate confidentiality and bandwidth efficiency in integral applications of IoT [3], [32].

2.2.2 Traditional Security Techniques

Access control, firewalls, and encryption are some of the traditional methods that offer partial protection. In practice, they fall short in IoT settings where devices are resource-constrained. These approaches also lack scalability and adaptability, which leads to performance issues under heavy traffic or in heterogeneous networks [9].

2.2.3 AI and Blockchain-Based Emerging Solutions

To address these gaps, recent studies have integrated artificial intelligence (AI) with blockchain. For example, [20] introduces a CNN-based method to optimize channel states in Industrial IoT, aiming to enhance efficiency and ensure secure data transmission. Similarly, [23] suggests a decentralized system involving the integration of Q-learning with reputation management to allow for anomaly detection and prioritize trusted devices. In 5G IoT scenarios, this architecture demonstrated improved intrusion detection and more equitable bandwidth allocation.

Other research explores both algorithms and system design. For example, [24] examines dynamic bandwidth and protocol selection. [25] proposes Blockchain-based Vehicle Identity Authentication (BVIA) and Dynamic Weighted Fair Bandwidth Allocation (DWFBA) for secure access and fair resource sharing in vehicular IoT. These methods effectively reduced task failure and communication overhead. In addition, [26] integrates deep reinforcement learning (DRL) with blockchain to improve consensus efficiency in permissioned networks, while [27] applies recurrent neural networks (RNNs) to predict storage failures and develop secure repair mechanisms to defend against eavesdropping.

Hybrid approaches have also been developed to improve quality of service (QoS). For instance, [28] applies an Analytic Hierarchy Process (AHP)-based matching game for fog computing optimization, while [29] introduces a trader metaheuristic algorithm for multi-objective bandwidth allocation. Predictive and adaptive designs such as Priority Dynamic Bandwidth Allocation (PDBA) [30] and freshness-aware frameworks [31] use reinforcement learning and Markov Decision Processes (MDPs) to reduce latency, lower packet loss, and maintain information freshness.

2.2.4 Outstanding Issues and Challenges

Despite these advances, several challenges remain High-complexity methods like DRL require extensive training data sets and high computational efforts, which could be impractical in bandwidth-limited IoT devices. Blockchain also has the latency and consensus overhead, and efficiency and security can be difficult to achieve in real-time systems. Developing an adaptive, lightweight, and scalable security scheme that is both dependable and equitable is an open problem in research.

2.2.5 Summary

Overall, security issues in bandwidth allocation have directed researchers toward AI, reinforcement learning, and blockchain-based trust systems. These technologies demonstrate clear improvements in intrusion detection, fairness, and resource utilization. However, challenges related to scalability, computational cost, and device heterogeneity still persist. Table 3 summarizes representative AI- and trust-based solutions for secure bandwidth allocation in IoT, highlighting their main contributions and performance improvements.

Table 3. Secure Bandwidth Allocation in IoT: AI and Trust-Based Techniques.

Paper	Year	Primary Focus	AI/ML Technique	Key Problem Solved	Performance Improvement
Gosw et al. [22]	2021	H-IoT Security/ Resources	CNN	Channel Security, Resource Utilization	Faster System, Better Resource Util.
Moudoud, H [23]	2023	5G+ IoT Security	DRL (Dist. Q-learning)	Intrusion Detection, Bandwidth Allocation	Outperforms referenced solutions
Bhar et al. [24]	2025	IoT Bandwidth	N/A (Survey/ Mechanisms)	BW Selection, Protocols, Topology	Enhanced communication stability
Liang et al. [25]	2023	Integrated IoT Security	Blockchain, DWFBA Algo.	Communication Security, Resource Fairness	↓ Comm. Overhead, ↓ Task Failures
Tsai et al. [26]	2022	IoT Blockchain	DRL	Blockchain Comm. Performance	Outperforms the current widely used algorithm.
Liao, C [27]	2019	IoT Data Reliability	RNN	Eavesdropping Prev., Repair BW	↑ 18.4% Security Level
Abedin, S et al. [28]	2019	Fog IoT Resources	AHP, Matching Game	User Assoc., Resource Allocation	↑ Utility Gain, Stable Association
Rouhifar, M [29]	2024	IoT Bandwidth	Trader Metaheuristic	Dynamic BW Distribution	↑ 6.32% Success Rate, ↑ 5.79% Throughput, ↑ 3.13% Resource Eff.
Chakour et al. [30]	2024	IoT Bandwidth	Predictive Algorithms (RL-inspired)	Min. Communication Delays	↓ 10% Latency, ↓ 6.8% Packet Loss, ↑ 94.2% Throughput, ↓ 0.69s Comp. Time
Guan, X et al. [31]	2023	Cellular IoT	MDP, Greedy Policy	Information Freshness (AoI)	Outperforms benchmarks

2.3 Heterogeneous IoT Network Management

Heterogeneous IoT (H-IoT) networks incorporate devices and applications with different protocols, hardware types, and demands. Heterogeneity introduces a set of management challenges. The most important ones include scalability, interoperability, energy efficiency, quality of service (QoS), and security. Conventional management frameworks struggle to effectively handle the dynamic and multi-faceted H-IoT in reality. As a result, unified and intelligent management solutions have become imperative [32].

2.3.1 Energy Efficiency and Anomaly Detection

To tackle abnormal data pattern and high energy consumption in H-IoT, [32] proposed an energy-efficient anomaly detection mechanism. The framework relies on the Parallelized Memetic Algorithm (PMA) with AlexNet. The Energy-Efficient Memetic Clustering Method (EEMCM) achieved over 99% accuracy for anomaly detection in IoT wireless sensor networks (WSNs). In practice, it improved scalability and robustness against abnormal traffic.

2.3.2 Hybrid Architectures for Industrial IoT

Industrial IoT (IIoT) environments demand diverse QoS guarantees, such as ultra-reliable low-latency communication (URLLC) and high data throughput. To address these requirements, [33,34] propose a hybrid RF/VLC architecture with reinforcement learning-based resource management. Using the Post-Decision State with Experience Replay and Transfer (PDS-ERT) algorithm, their framework enhances both energy efficiency and QoS in heterogeneous smart factory networks.

2.3.3 Privacy and Multi-Mode Networks

In multi-mode H-IoT settings, privacy preservation and queuing delays remain critical concerns. [35] presents a privacy-aware optimization framework that leverages Lyapunov optimization and auction-based matching to jointly minimize latency and maximize privacy entropy. Similarly, [36] introduces a Software-Defined Networking (SDN)-based IoT management scheme that homogenizes heterogeneous controller response times, thereby co-optimizing QoS and security in highly diverse IoT systems.

2.3.4 Interoperability and Large-Scale Coordination

Achieving interoperability across large-scale heterogeneous IoT systems requires advanced orchestration mechanisms. In [37], large-scale HetNet simulations are used to compare centralized and distributed management approaches, demonstrating trade-offs between scalability, reliability, and packet loss. To address decentralization, [38] proposes a blockchain-based architecture for secure and transparent coordination in large-scale H-IoT, reducing single points of failure and enabling trusted collaboration.

2.3.5 Edge and Federated Management Frameworks

Edge computing and federated approaches offer additional mechanisms for handling heterogeneity. [39] formulates a mixed-integer program for Multi-Access Edge Computing (MEC) resource provisioning and workload assignment, optimizing trade-offs between cost, latency, and QoS. In federated contexts, [40] introduces Hetero-FedIoT, a rule-based interworking framework that supports cross-platform interoperability and adaptive aggregation to improve model convergence across diverse IoT environments.

2.3.6 AI/ML-Driven Management Solutions

Artificial Intelligence and Machine Learning (AI/ML) techniques have become central to addressing heterogeneity and resource management challenges in IoT. Comprehensive surveys such as [41,42] emphasize the potential of ML/DL for anomaly detection, resource optimization, and interoperability enhancement, offering a roadmap for intelligent, large-scale IoT management frameworks.

2.3.7 Summary

In conclusion, managing heterogeneous IoT networks requires integrated solutions that simultaneously address QoS, security, interoperability, and energy efficiency. Existing contributions span from anomaly detection frameworks and hybrid RF/VLC architectures to SDN-enabled management, blockchain-based coordination, and federated learning approaches. AI and ML remain at the core of these advancements, enabling predictive and adaptive management for highly dynamic IoT ecosystems. Table 4 provides a consolidated summary of AI/ML-driven methods for H-IoT management and their reported performance improvements.

Table 4. AI/ML-Driven Solutions for Managing Heterogeneous IoT Networks.

Paper	Year	Primary Focus	AI/ML Technique	Key Problem Solved	Performance Improvement
Thangavel, A [32]	2024	H-IoT Energy/Security	Memetic Algo. + AlexNet CNN	Anomaly Detection, Energy Eff.	99.11% Anomaly Detection Accuracy
Yang, H et al. [33]	2020	Smart Factory QoS	Deep PDS-ERT RL	Energy-Eff. Resource Mgmt (RF/VLC)	Superior perf. (energy eff. & QoS)
Gan, Z et al. [34]	2025	Smart Factory QoS	Deep PDS-ERT RL	Energy-Eff. Resource Mgmt (RF/VLC)	Superior perf. (energy eff. & QoS)
Sood, K et al. [35]	2020	Multi-mode H-IoT	Lyapunov Opt., Auction Matching	Queuing Delay, Privacy, Security	↓ Queuing Delay, ↑ Privacy, Security

Selvara, S et al. [36]	2021	SDN-IoT QoS/Security	Homogenization Framework	Heterogeneity in SDN-IoT (QoS/Security)	Significantly alleviates heterogeneity
Tsenga, L et al. [37]	2020	HetNets Interoperability	Simulation Model	Latency, Reliability in HetNets	↓ 100-fold Mean Latency (centralized vs. distributed)
Kherf, N et al. [38]	2019	H-IoT Management	Blockchain (Architecture)	Managing Large-scale H-IoT	N/A (Preliminary study)
Khan, A et al. [39]	2024	MEC Resource Provisioning	Decomposition Approach	MEC Resource & Workload Assignment	Highlights performance trends/trade-offs
Zafar, A et al. [40]	2023	Federated H-IoT	Rule-based Interworking, Aggregation	Seamless Connectivity, Interoperability	Superior comp. & comm. efficiency
Husain, F et al. [41]	2019	SDN-IoT QoS	Homogenization Framework	Heterogeneity in SDN-IoT QoS	Some improvement (alleviates heterogeneity, maintains QoS)

2.4 Literature Synthesis and Research Directions

From the reviewed literature in Tables 1–3, key trends and gaps are evident. Advances have been made in three areas: dynamic bandwidth allocation, secure bandwidth management, and heterogeneous IoT optimization. Dynamic allocation has shifted from static methods to AI-driven techniques such as deep learning and reinforcement learning, improving latency, throughput, and utilization. Security enhancements include blockchain, reputation systems, and AI-based intrusion detection, while heterogeneous IoT management has leveraged hybrid architectures, federated learning, and blockchain for better scalability, energy efficiency, and QoS.

Despite this progress, major limitation persists: (1) absence of unified frameworks integrating allocation, security, and heterogeneity; (2) limited large-scale, real-time validation; (3) partial integration of AI and blockchain; and (4) low adaptability to emerging technologies like 6G and Edge AI.

This study addresses these limitations by proposing a Q-learning and blockchain-enabled framework for secure, dynamic bandwidth allocation in heterogeneous IoT networks. A proof-of-concept demonstrates trust-based decision-making integration, with a modular design adaptable to various protocols and future technologies such as SDN and federated learning. Thereby laying a foundation for scalable, intelligent, and secure IoT infrastructures.

3. Background

The continuous growth of the Internet of Things (IoT) has reshaped network requirements, especially in terms of security, interoperability, and scalability. With the increasing diversity of IoT devices and protocols, reliable performance can only be achieved through effective resource control, fair bandwidth distribution, and secure communication mechanisms.

To address these concerns, this section presents the background of our framework by discussing five essential aspects. The first is the heterogeneous nature of IoT, which reflects the wide variety of devices and communication technologies. The second is Q-learning, an adaptive reinforcement learning approach for resource allocation. The third is blockchain, which serves as a decentralized solution for trust and security. The fourth concerns IoT trust models that help distinguish trustworthy devices from malicious ones. The fifth involves the mathematical formulation that underpins the entire system. Together, these components provide the conceptual and technical basis for the integrated blockchain-supported Q-learning model proposed in this work.

3.1 Synthesis of Literature and Research Directions

IoT is not all the same. It brings together devices that use different protocols, have different processing power, and even ask for different amounts of energy. This happens because many technologies live side by side, like Wi-Fi, Bluetooth, Zigbee, LoRaWAN, and also the cellular networks. When all these mix, handling resources and keeping things compatible gets messy [36].

There is also the matter of trust. Some devices are strong and secure, like sensors used in industry. Others are very simple, like home gadgets. This big gap makes it harder to keep the system reliable and to make sure the quality of service (QoS) is good enough [37].

To deal with such variety, we need architectures that can balance things like energy, latency, scalability, and security, while still letting devices work smoothly together. Some studies talked about hybrid or rule-based architectures as ways to connect devices in this mixed IoT world [39]. These give more freedom and make management easier when the system grows. Other works looked at edge computing and software-defined networking (SDN). These tools help to manage device diversity with real-time changes and better QoS in different IoT environments [35], [40].

3.2 Q-learning Basics

Q-learning is an RL model-free algorithm. It enables agents to learn optimal action by interacting with the environment without knowing the behavior of the system beforehand [13]. Essentially, Q-learning adjusts a Q-value for the expected reward of performing action in a state and then pursuing the optimal policy [14].

The algorithm works by learning the world, receiving feedback in the form of reward or penalty, and updating the Q-table according to the Bellman equation. Q-values converge to virtually optimal values over a period of time, and the agent chooses actions to achieve maximum long-term performance [14]. This makes Q-learning perfect for dynamic IoT scenarios where bandwidth requests, delay, and trust are dynamically changing. Research has demonstrated its ability in resource allocation, intrusion detection, and energy reduction in IoT, exhibiting enhanced flexibility and robustness compared to rule-based methods [23], [24].

3.3 Blockchain Fundamentals

Blockchain is a distributed ledger technology that offers transparency, security, and immutability [25]. All transactions were inserted into a block, which was cryptographically linked to the previous block, thus creating an unalterable chain. The system is distributed and utilized consensus protocols such as Proof-of-Work (PoW) or Proof-of-Stake (PoS) for validation [26].

Blockchain is increasingly recognized as an effective mechanism for establishing trust and security among heterogeneous IoT devices. It provides secure data sharing, authentication, and resource delegation even in untrusted environments [22]. Through its decentralized trust layer, blockchain eliminates single points of failure and mitigates attacks such as data tampering and bandwidth hijacking [25]. Recent research has also explored the integration of blockchain with AI and SDN to enable more adaptive and scalable IoT systems. For instance, in blockchain-based resource allocation models, smart contracts are employed to autonomously manage bandwidth distribution and enforce trust-based decisions, thereby ensuring efficiency, security, and fairness [26].

3.4 Trust Models in IoT

Trust management is a critical component of heterogeneous IoT, where devices differ in terms of reliability, security capabilities, and behavior. Unlike traditional networks, IoT environments often include resource-constrained or even malicious devices. This makes trust assessment essential for enabling secure communication and ensuring fair bandwidth allocation [22], [25].

Trust models generally analyze devices using multiple parameters such as historical behavior, packet delivery ratio, latency, and energy usage. By differentiating between trustworthy and untrustworthy nodes, these models enhance QoS and overall network resilience [27]. Reputation systems, for instance, assign trust scores to devices based on their interaction history. Blockchain-inspired trust models add an immutable layer of verification for device actions and transaction activities [25], [26]. Recent studies further combine trust evaluation with machine learning and reinforcement learning, where Q-learning agents dynamically adapt bandwidth allocation according to trust ratings [23]. In such approaches, malicious or low-trust devices are penalized, while trusted devices receive equitable and efficient bandwidth assignments.

Hybrid trust models employ multi-objective optimization techniques, such as AHP-based allocation for fog computing, to achieve fairness, efficiency, and flexibility [28]. Integrating trust management with AI and blockchain significantly strengthens both security and adaptability in complex IoT environments.

3.5 Mathematical Formulation of the Framework

To formalize the proposed model, we define the following equations governing trust dynamics, admission control, learning updates, and performance metrics.

Equation (1): Trust Update

The trust value of each device is dynamically updated depending on whether it behaves legitimately or maliciously:

$$Ti(\mathrm{t}+1) = \begin{matrix} \min(1, Ti(\mathrm{t}) + \Delta^{+}), \text{if device i is legitimate} \\ \max(0, Ti(\mathrm{t}) - \Delta^{-}), \text{if device i is malicious} \end{matrix} \qquad (1)$$

Where $Ti(\mathrm{t})$ Trust value of device i at time step t, range [0,1]. Δ^{+} Increment step for legitimate devices. Δ^{-} Decrement step for malicious devices.

Equation (2): Service Admission Condition

Only devices whose trust score exceeds the threshold are admitted for service:

$$Service(i) = \begin{matrix} 1, if\ Ti \geq \theta\ and\ device\ i\ is\ non\ attacker \\ 0\ , \qquad\qquad \text{otherwise} \end{matrix} \qquad (2)$$

Where Ti is the Trust score of devices ***i***. θ represents Trust threshold. $Service(i)$ represents binary admission variable (1 = admitted, 0 = rejected).

Equation (3): Q-learning Update Rule

The reinforcement learning agent updates its policy using the Bellman equation:

$$\mathrm{Q}(\mathrm{s},a) \leftarrow \mathrm{Q}(\mathrm{s},a) + \alpha\,[\,\mathrm{r} + \gamma\, max_{\mathrm{a}}{}'\,\mathrm{Q}(\mathrm{s}',a') - \mathrm{Q}(\mathrm{s},a)\,] \qquad (3)$$

Where $\mathrm{Q}(\mathrm{s},\mathrm{a})$ is the Q-value for state s and action a, α is the learning rate, γ is the discount factor, r is the reward received, s′ represents next state, $\max_{\mathrm{a}}{}'\ \mathrm{Q}(\mathrm{s}',a)$ represents Maximum expected Q-value of the next state.

Equation (4): Served Ratio (SR)
This metric measures the proportion of devices that successfully receive service:

$$\mathrm{SR} = \frac{1}{N}\sum_{i=1}^{N}\begin{cases}1 & if\ bw_i > 0\\ 0 & \text{otherwise}\end{cases} \qquad (4)$$

Where N is the total number of devices, bw_i represents bandwidth allocated to device i, SR represents served ratio.

Equation (5): Average Bandwidth
The mean allocated bandwidth per device is given as:

$$\mathrm{AvgBW} = \frac{1}{N}\sum_{i=1}^{N} bw_i \qquad (5)$$

Where AvgBW represents bandwidth per device.

Equation (6): Jain's Fairness Index
To evaluate fairness in bandwidth allocation, Jain's index is computed as:

$$J = \frac{(\sum_{i=1}^{N} bw_i)^2}{N\sum_{i=1}^{N}(bw_i{}^2) + \delta} \qquad (6)$$

Where J represents jain's fairness index, ranges between 0 and 1. δ represents small constant to prevent division by zero.

Equation (7): Energy Consumption
The total communication energy consumption is expressed as:

$$E_{total} = \sum_{i=1}^{N}(bw_i \times 1000 \times E_{bit}) \qquad (7)$$

Where E_{total} is the total energy consumption across devices, E_{bit} represents energy consumed per bit.

4. Proposed Framework (Architecture)

4.1 Overall Description

The proposed architecture facilitates secure and dynamic bandwidth allocation in heterogeneous IoT networks. It integrates three main components: (1) an adaptive Q-learning engine that optimizes resources in real time, (2) a blockchain-based trust layer that enforces security and fairness, and (3) a feedback-driven assessment mechanism that enables continuous improvement.

By incorporating these elements, the framework ensures that trustworthy devices receive sufficient resources, while malicious or wasteful nodes are constrained. This design directly addresses critical challenges such as device heterogeneity, latency, energy constraints, and evolving security threats.

Figure 1. High-level overview of the proposed unified framework, highlighting the integration of machine learning-based dynamic bandwidth allocation, blockchain-enabled security enhancement, and adaptive handling of heterogeneous IoT devices to achieve secure and real-time network optimization.

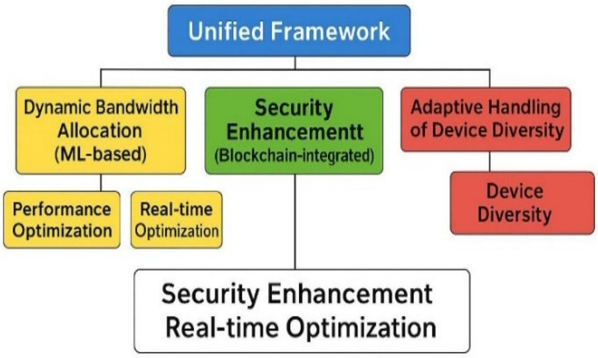
Unified Framework
Dynamic Bandwidth Allocation (ML-based)
Security Enhancementt (Blockchain-integrated)
Adaptive Handling of Device Diversity
Performance Optimization
Real-time Optimization
Device Diversity
Security Enhancement Real-time Optimization

Fig. 1. High-level overview of the proposed unified framework

4.2 Three-Layer Framework Structure

The proposed framework adopts a three-layer architectural design to ensure modularity, scalability, and clear functional separation among system components, making it suitable for large-scale heterogeneous IoT environments.

(A) Data Collection Layer aggregates bandwidth demands, latency measurements, and trust values from a wide range of IoT devices.

(B) Data Processing Layer (Q-learning) analyzes these inputs and dynamically allocates bandwidth to optimize both efficiency and fairness.

(C) Data Leverage Layer (Blockchain Trust) validates allocation decisions, maintains trust scores, and disseminates information in a secure and transparent manner.

This layered design preserves modularity and scalability, making it well-suited for large-scale IoT deployments.

4.3 System Workflow

The execution workflow of the proposed framework is illustrated in Fig. 2, which depicts the sequential interaction among the learning, security, and network components.

The process begins when IoT devices generate bandwidth requests accompanied by their current trust-related attributes. The aggregated network state is then observed by the Q-learning engine, which selects an appropriate bandwidth allocation action based on its learned policy.

Subsequently, the proposed allocation decision is forwarded to the blockchain-based trust layer for validation. Only devices that satisfy the trust requirements are admitted and granted bandwidth access, while untrusted or malicious nodes are blocked or denied service.

Following the admission decision, the system performance is evaluated in terms of latency, fairness, and energy efficiency. The resulting performance feedback is then used to update both the Q-learning policy and device trust states, forming a temporal learning loop that continuously refines bandwidth allocation decisions across successive time steps.

This workflow ensures that bandwidth allocation is not only adaptive and performance-aware, but also securely enforced through trust validation and feedback-driven learning.

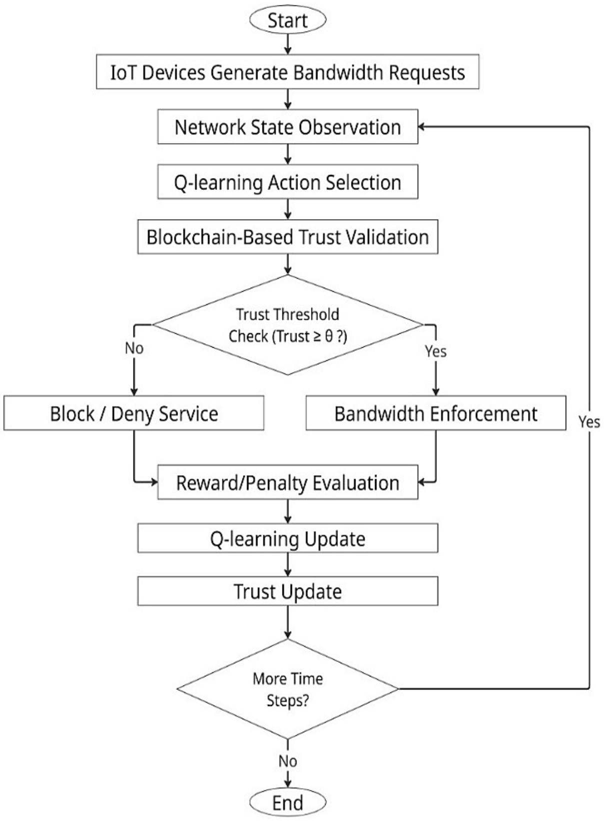
Start
IoT Devices Generate Bandwidth Requests
Network State Observation
Q-learning Action Selection
Blockchain-Based Trust Validation
Trust Threshold Check (Trust ≥ θ ?)
No
Yes
Block / Deny Service
Bandwidth Enforcement
Yes
Reward/Penalty Evaluation
Q-learning Update
Trust Update
More Time Steps?
No
End

Fig. 2. System Workflow of the Proposed Hybrid Q-Learning and Blockchain Framework.

4.4 Blockchain Model Specification and Overhead Analysis

A lightweight, permissioned blockchain model is A lightweight, permissioned blockchain model is integrated at the edge layer to enable trust-aware bandwidth allocation in heterogeneous IoT networks. Resource-intensive public or mining-based blockchains are avoided due to their excessive latency and energy overhead, while permissioned, edge-assisted blockchain architectures are more suitable for IoT environments [8], [9], [37]. In the proposed framework, IoT devices do not directly participate in blockchain operations; instead, trusted edge gateways act as authority nodes responsible for ledger maintenance and transaction validation under a Proof-of-Authority (PoA)–based setting [10], [26].

The blockchain ledger stores only decision-level metadata rather than raw sensing or traffic data. Each transaction corresponds to a validated bandwidth allocation event and includes ⟨Node ID, trust score, allocated bandwidth, timestamp⟩, following lightweight blockchain-based trust management designs for IoT systems [8], [25]. Transactions are generated in an event-driven manner and recorded only when a device satisfies the trust threshold and is granted bandwidth. For fair comparison, the blockchain layer is activated exclusively in the Blockchain-only and Hybrid scenarios.

The blockchain overhead is analytically characterized in terms of latency, energy consumption, and communication cost. Let N_{tx}denote the number of blockchain transactions and $S_{tx}$the transaction size in bits. The blockchain-related energy consumption is modeled as

$$E_{bc} = N_{tx} \times S_{tx} \times E_{bit} \tag{8}$$

where E_{bit}represents the energy consumed per transmitted bit. The communication overhead is expressed as

$$C_{bc} = N_{tx} \times S_{tx} \tag{9}$$

In addition, the blockchain-induced latency is modeled as

$$L_{bc} = N_{tx} \times L_{tx} \tag{10}$$

where L_{tx}denotes the per-transaction validation and confirmation delay under a lightweight permissioned consensus. These formulations enable systematic quantification of blockchain overhead while preserving a practical and deployment-oriented design for heterogeneous IoT networks [10], [28].

4.5 Simulation Setup

The proposed framework was evaluated using a MATLAB-based simulation conducted on a heterogeneous IoT network comprising 100 devices. Each device was randomly assigned a communication protocol (e.g., ZigBee, Wi-Fi, LoRa, or NB-IoT), along with distinct latency and initial trust values to reflect realistic network diversity. A subset of devices was configured to exhibit malicious behavior, allowing the assessment of trust dynamics and security resilience under adversarial conditions.

Dynamic bandwidth allocation was performed using a Q-learning agent operating over ten discrete bandwidth levels within a predefined range. Trust scores were continuously updated through the blockchain-based trust validation mechanism described in Section 4.4, ensuring that allocation decisions were influenced by both performance feedback and security considerations.

The simulation was executed over 100 time steps, during which multiple allocation strategies including static allocation, AI-only, blockchain-only, hybrid learning-based allocation, and classical scheduling baselines were evaluated under identical conditions. Key performance metrics such as average allocated bandwidth, served ratio, Jain's fairness index, total energy consumption, and security-related indicators were collected and analyzed.

The results were visualized using time-series plots, histograms, and comparative bar charts to examine convergence behavior, resource fairness, energy efficiency, and the impact of trust-based admission control.

Table 5. Simulation Parameters.

Parameter	**Value / Description**
Number of IoT devices (N)	100
Simulation time (T)	100 time steps
Bandwidth range (bw_i)	50 – 500 kbps (10 discrete levels).
Trust threshold (θ)	0.6
Learning rate (α)	0.5
Discount factor (γ)	0.9
Energy per bit (E_{bit})	1×10^{-6} J/bit
Blockchain transaction size S_{tx}	256 bits
Blockchain transactions N_{tx}	Event-driven (trust ≥ θ and bw > 0)
Blockchain latency per transaction L_{tx}	5 ms
Performance metrics	Average Bandwidth, Served Ratio, Jain's Fairness, Total Energy, Trust Accuracy, TPR, FPR

5. Results and Discussion

This section evaluates the performance of the proposed framework under various operational scenarios and discusses the observed behaviors in light of the design choices introduced in Section 4. Rather than reporting raw numerical outcomes in isolation, the discussion focuses on understanding why certain strategies succeed, where trade-offs emerge, and how security-aware learning reshapes bandwidth allocation decisions over time. To ensure a fair and comprehensive assessment, the proposed hybrid framework is compared against multiple baseline strategies, including static allocation, AI-only learning, blockchain-only trust enforcement, classical scheduling algorithms (WFQ and DRR), and a deep reinforcement learning baseline (DQN). All scenarios are evaluated under identical network conditions and simulation parameters, allowing performance differences to be attributed directly to the underlying allocation and security mechanisms.

5.1 Overall Performance Comparison

Figure 3 presents an overall comparison of the four core bandwidth allocation strategies Static, AI-only, Blockchain-only, and the proposed Hybrid framework across average bandwidth, energy consumption, served ratio, and trust accuracy.

The Static strategy achieves the highest bandwidth utilization and full-service coverage due to uniform allocation; however, this behavior results in excessive energy consumption and does not consider device reliability.

The AI-only approach introduces adaptive bandwidth adjustment, leading to moderate reductions in bandwidth usage and energy consumption. Nevertheless, the absence of trust awareness allows unreliable devices to continue consuming resources, which negatively affects fairness and security-related performance indicators.

In contrast, the Blockchain-only strategy prioritizes trust enforcement by filtering low-trust devices, resulting in the lowest energy consumption among the evaluated schemes and improved trust accuracy. This strict security-oriented behavior, however, leads to conservative bandwidth allocation and a reduced served ratio.

The proposed Hybrid framework balances these trade-offs by integrating Q-learning–based adaptation with blockchain-inspired trust validation. As shown in Fig. 3, the Hybrid model maintains controlled bandwidth utilization and improved energy efficiency while achieving the highest trust accuracy. Moreover, it attains a served ratio superior to the Blockchain-only scheme, demonstrating the benefit of combining learning-driven adaptability with trust-aware filtering for secure and efficient bandwidth management in heterogeneous IoT environments.

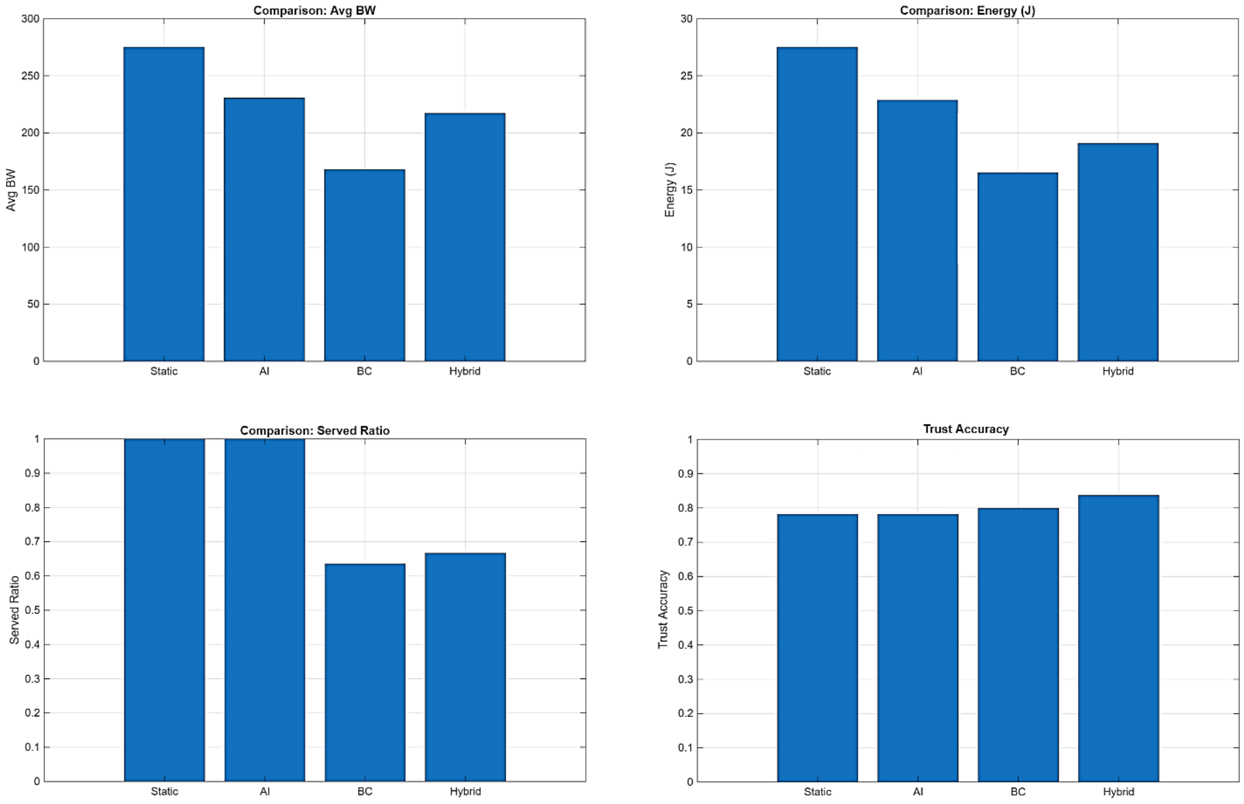

Fig. 3: Overall performance comparison of the four core allocation strategies in terms of bandwidth utilization, service availability, fairness, and energy consumption.

The aggregated performance metrics of the four core strategies are summarized in Table 6, complementing the comparative trends discussed above.

Table 6. Comparative Performance Metrics of Core Bandwidth Allocation Strategies (Static, AI-only, Blockchain-only, and Hybrid).

Scenario	AvgBW (kbps)	Served Ratio	Jain's Fairness	Total Energy (J)	Trust Accuracy	TPR	FPR
Static	275	1	1	27.5	0.79	0	0
AI-only	240	1	0.52	23.8	0.79	0	0
Blockchain-only	170	0.62	0.6	16.8	0.8	1	0.23
Hybrid (Q+Trust)	**220**	**0.68**	**0.66**	**19.5**	**0.83**	**1**	**0.20**

5.2 Analysis of Hybrid Framework Behavior

This subsection analyzes the temporal behavior and convergence characteristics of the proposed hybrid framework by examining the evolution of average bandwidth, energy consumption, served ratio, and trust distribution over time.

As shown in Figure 4, the average allocated bandwidth increases rapidly during the initial learning phase, reflecting the exploratory behavior of the Q-learning agent. After approximately 25–30 time steps, the allocation stabilizes within a consistent operating range, indicating convergence toward a steady policy shaped by trust constraints.

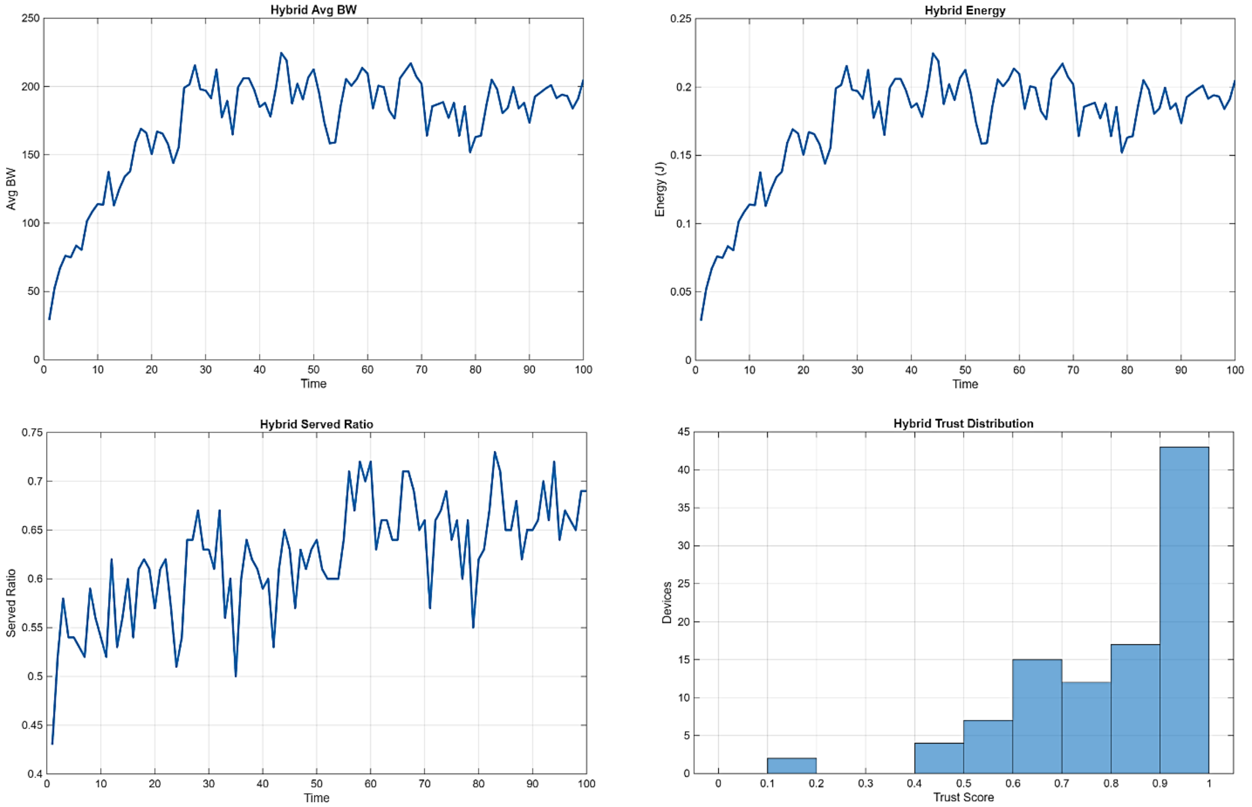

Fig. 4: Temporal behavior and convergence characteristics of the proposed hybrid framework

Energy consumption follows a similar trajectory, with early fluctuations gradually diminishing as learning converges. This behavior confirms that the hybrid framework internalizes longer-term traffic and trust patterns rather than reacting solely to instantaneous conditions, resulting in stable and energy-efficient operation.

The served ratio improves progressively as trustworthy devices are increasingly prioritized, while temporary variations reflect dynamic traffic demands and trust-based filtering. The final trust distribution demonstrates that most devices converge toward high trust scores, whereas low-trust or malicious nodes are effectively isolated.

Overall, these results indicate that the hybrid framework converges smoothly without oscillatory instability or excessive conservatism. The interaction between learning-based adaptation and trust validation remains cooperative rather than brittle, enabling stable, adaptive, and security-aware bandwidth allocation in heterogeneous IoT environments.

5.3 Comparison with Classical Scheduling Algorithms (WFQ and DRR)

Classical scheduling algorithms such as WFQ and DRR rely on predefined service rules and do not incorporate learning or trust awareness. As shown in Figure 5, both algorithms achieve high fairness indices and stable service availability by distributing bandwidth according to fixed scheduling principles. WFQ provides proportional fairness based on assigned weights, while DRR ensures starvation-free service through deterministic rotation.

However, this strict fairness comes at the cost of contextual awareness. Neither WFQ nor DRR differentiates between legitimate and malicious devices, treating all nodes equally regardless of behavior. Under adversarial conditions, such fairness becomes indiscriminate, allowing unreliable devices to consume resources without penalty.

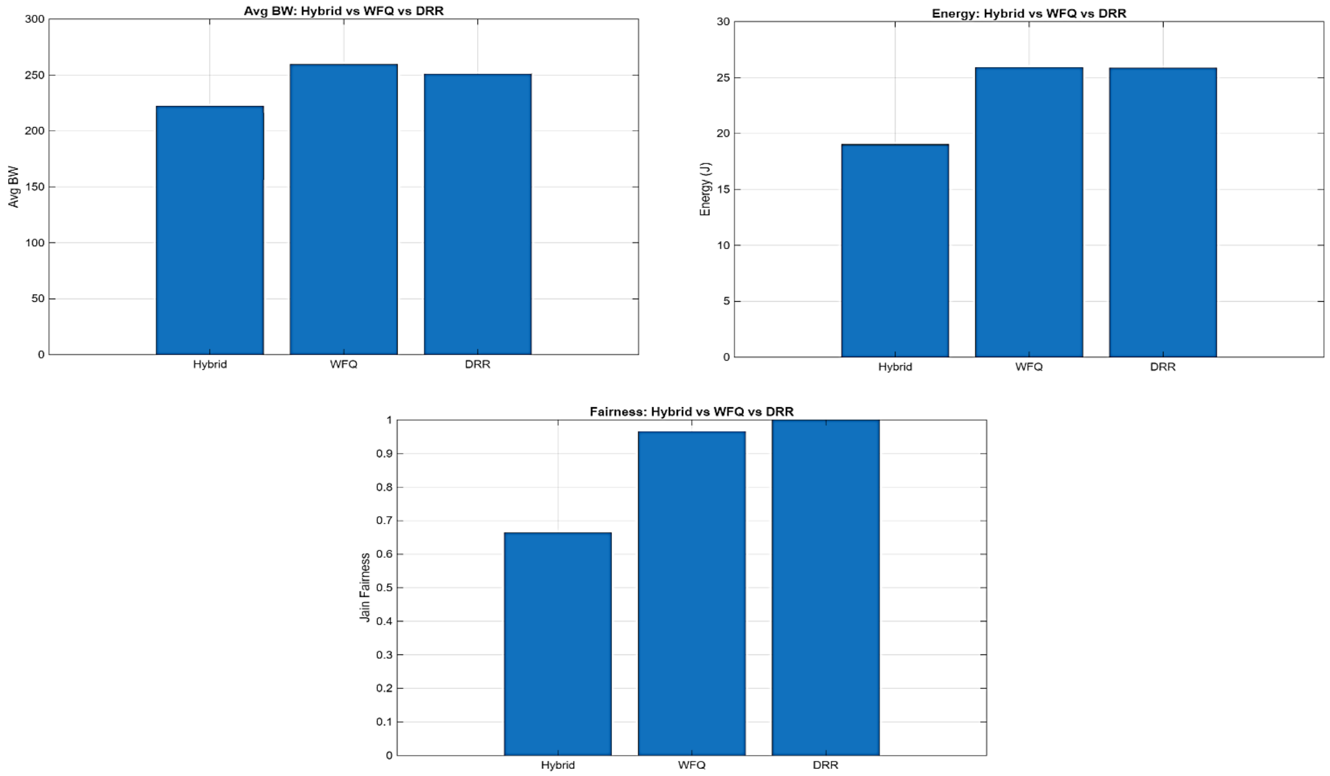

Fig. 5: Performance comparison between the proposed hybrid framework and classical scheduling algorithms (WFQ and DRR)

Energy consumption further highlights this limitation. Because classical schedulers continue serving all devices uniformly, they maintain higher energy usage even when transmissions are inefficient or harmful. In contrast, the proposed hybrid framework selectively allocates resources based on both learned network dynamics and trust validation. While the hybrid approach does not outperform WFQ or DRR in raw fairness under ideal cooperative conditions, it achieves superior robustness and energy efficiency in heterogeneous and partially adversarial environments. By coupling adaptive learning with trust-based filtering, the hybrid framework prioritizes meaningful service delivery over blind fairness, making it more suitable for realistic IoT deployments where cooperation cannot be assumed.

5.4 Hybrid Framework versus Deep Reinforcement Learning (DQN)

Figure X compares the proposed hybrid framework with a Deep Q-Network (DQN) in terms of average bandwidth, energy consumption, and served ratio. As expected, DQN achieves higher average bandwidth and service availability due to its expressive function approximation and aggressive learning strategy. By leveraging deep neural networks, DQN rapidly adapts to network dynamics and optimizes throughput-oriented objectives.

However, these performance gains are accompanied by increased energy consumption and higher learning overhead. The absence of explicit trust awareness causes DQN to treat all devices uniformly, allowing unreliable or malicious nodes to influence allocation decisions. As a result, optimization focuses on observable performance metrics rather than behavioral reliability.

In contrast, the proposed hybrid framework adopts a constrained learning strategy. Instead of maximizing throughput aggressively, Q-learning operates within trust-based validation boundaries enforced by the blockchain-inspired layer. This design leads to slightly lower average bandwidth and served ratio compared to DQN, but significantly improves control, stability, and energy efficiency.

The hybrid framework also exhibits reduced oscillations and faster stabilization under limited training horizons. Its lower computational complexity, absence of deep neural models, and embedded security awareness make it more suitable for resource-constrained and security-sensitive IoT environments.

Overall, while DQN represents a high-performance solution under unconstrained optimization, the hybrid framework offers a more practical and robust alternative for heterogeneous IoT deployments, where energy efficiency, trust enforcement, and predictable behavior are as critical as raw throughput.

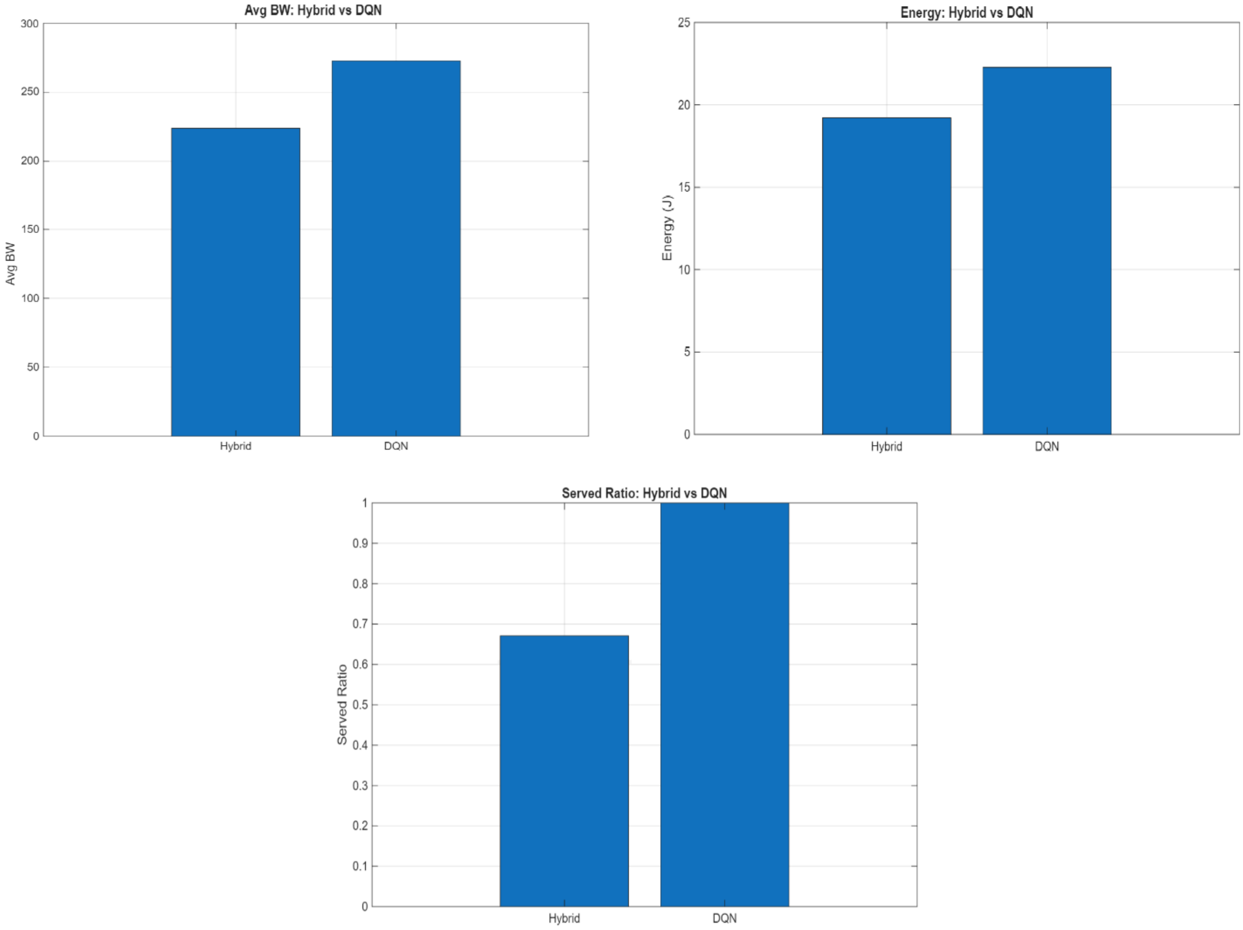

Fig. 6: Performance comparison between the proposed hybrid framework and the DQN-based allocation approach.

5.5 Security and Trust Evaluation

Security in the proposed framework is modeled as a dynamic, evidence-driven process rather than a binary decision. Trust values evolve gradually based on observed device behavior, enabling reliable differentiation between benign and malicious nodes without overreacting to transient anomalies. This approach reduces sensitivity to noise while preserving consistent behavioral assessment.

The blockchain-inspired trust layer maintains historical accountability by recording validated allocation decisions, allowing trust to be updated instead of being permanently fixed. Consequently, malicious behavior is detected reliably once sufficient evidence accumulates, while occasional false positives remain unavoidable in realistic IoT environments. Under the evaluated conditions, this is reflected by complete attack detection (TPR = 1.0) alongside a non-zero false-positive rate, indicating a practical balance between sensitivity and robustness rather than idealized classification.

Unlike rigid security-only enforcement, the hybrid framework preserves adaptability under trust constraints. Learning decisions are filtered rather than overridden, allowing legitimate devices to recover from temporary trust degradation, while persistent malicious activity leads to progressive bandwidth restriction without abrupt exclusion.

Security is embedded directly into the learning loop rather than applied as an external control mechanism. Trust validation constrains the action space explored by the Q-learning agent, shaping long-term policy formation instead of post-hoc enforcement. This integration enables adaptive bandwidth allocation while containing adversarial impact.

To further characterize the security layer, the blockchain-related overhead is quantified for blockchain-enabled scenarios. Table 7 summarizes the associated latency, energy consumption, and communication cost based on the event-driven transaction model.

Table **7**. Blockchain Overhead under Security and Trust Evaluation

Scenario	**Latency (ms)**	**Energy (J)**	**Communication Cost (bits)**
Blockchain-only	5	0.0044	4.4×10^4

Hybrid (Q + Trust)	5	0.0049	4.9×10^4

Overall, the results indicate that blockchain overhead remains limited due to event-driven transaction generation and lightweight validation. The hybrid framework incurs only a marginal overhead increase compared to the blockchain-only scheme, while providing improved adaptability and resilience, confirming the feasibility of integrating trust-aware security into learning-based bandwidth allocation.

5.6 Trust-Aware Learning-Based IoT Frameworks: Discussion

Recent studies published in the EAI Endorsed Transactions on Internet of Things highlight the growing interest in trust-aware learning mechanisms for securing IoT systems, particularly in critical and privacy-sensitive environments [43]. These works primarily focus on enhancing model robustness, privacy preservation, or secure collaboration, often assuming that resource management operates independently of trust enforcement.

In contrast, the framework proposed in this study integrates trust directly into the bandwidth allocation process. Trust validation influences the learning decisions themselves rather than acting as an external or post-processing constraint. This design allows adaptive allocation to remain responsive while being progressively shaped by observed device behavior.

Unlike trust-aware federated learning approaches that emphasize global model convergence, the proposed framework targets localized, real-time decision-making under heterogeneous traffic and uncertain conditions. By embedding trust within the learning loop, the system supports controlled adaptation without relying on rigid exclusion or static security rules.

Overall, this perspective complements existing trust-aware IoT studies by addressing trust as a governing factor in dynamic resource allocation, rather than as a mechanism limited to data or model protection. This positioning clarifies the contribution of the proposed approach within the broader landscape of trust-aware learning-based IoT frameworks.

6. Conclusion

This work addressed a challenge that rarely appears in isolation. In heterogeneous IoT networks, bandwidth efficiency, security, and adaptability are often treated as separate objectives, optimized independently and reconciled later, if at all. The proposed framework follows a different approach by modeling bandwidth allocation as a decision-making process jointly shaped by learning and trust, rather than as a static rule or an unconstrained optimization task.

By integrating Q-learning with a lightweight, permissioned blockchain-based trust layer, the framework demonstrates that adaptive bandwidth allocation does not need to be blind, and security enforcement does not need to be rigid. Learning provides flexibility, trust provides restraint, and their interaction yields a system that evolves rather than oscillates. Across multiple scenarios, the observed behavior indicates that meaningful performance improvements arise not from extreme optimization, but from balance.

The comparative evaluation reinforces this perspective. Static allocation remains predictable but inflexible, while learning-only approaches adapt quickly yet remain vulnerable to exploitation. Blockchain-only enforcement strengthens security at the cost of responsiveness, whereas classical schedulers preserve fairness without contextual judgment. Deep reinforcement learning further improves performance, often at the expense of control and stability. Rather than dominating these approaches individually, the proposed hybrid framework aligns their strengths while mitigating their weaknesses. Its evaluation against both classical schedulers (WFQ and DRR) and deep reinforcement learning baselines highlights the distinction between unconstrained optimization and trust-aware, deployment-oriented decision-making—particularly under adversarial conditions where resilience outweighs peak throughput.

Equally important are the claims the framework does not make. It does not assume instant attack elimination or perfect trust classification, which are unrealistic in dynamic IoT environments. Instead, threats are contained progressively, decisions are revised when necessary, and stability is maintained in the presence of noise and uncertainty. Trust is earned, lost, and occasionally regained, while allocation decisions remain open to correction. This behavior reflects practical deployment realities rather than idealized assumptions.

Several directions for future work emerge from this study. Extending the framework toward federated or multi-agent learning could enhance scalability in large deployments. Incorporating software-defined networking may further improve enforcement granularity, while real-world prototyping could validate the abstraction choices made at the edge layer. These extensions build naturally on the foundation established here.

In conclusion, secure and dynamic bandwidth allocation in heterogeneous IoT networks is not a matter of choosing between learning and security, but of designing how they coexist. When learning is constrained by trust and trust is informed by experience, the result is a system that behaves less like a rigid algorithm and more like a well-governed network.

Author Contributions

Ahmed Saeed: Conceived the research idea, conducted a comprehensive literature review, designed the methodology, implemented the simulation software, analyzed the data, and prepared the initial draft of the manuscript.

Ahmed Shahidan: Provided academic supervision, validated the research design, and conducted critical review and editing of the manuscript.

Hind Taha: Contributed to the development of the simulation framework, assisted in interpreting and analyzing the results, and provided technical insights for enhancing heterogeneous IoT networks.

Sanaa Tarish: Organized the data, prepared and formatted the tables, and arranged the key references and sources for the manuscript.
All authors read and approved the final manuscript.

Funding Declaration

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

References

[1] Alatram, A., Sikos, L. F., Johnstone, M., Szewczyk, P., & Kang, J. J., "DoS/DDoS-MQTT-IoT: A dataset for evaluating intrusions in IoT networks using the MQTT protocol," *Computer Networks*, vol. 231, p. 109809, 2023. https://doi.org/10.1016/j.comnet.2023.109809

[2] Hong, J., Hong, Y-G., de Foy, X., Kovatsch, M., Schooler, E., and D. Kutscher, "Internet of Things (IoT) Edge Challenges and Functions", RFC 9556 Internet Research Task Force (IRTF), April 2024, https://www.rfc-editor.org/info/rfc9556.

[3] Bukhowah, R., Aljughaiman, A., & Rahman, M. H. (2024). Detection of dos attacks for IoT in information-centric networks using machine learning: Opportunities, challenges, and future research directions. Electronics, 13(6), 1031. https://doi.org/10.3390/electronics13061031

[4] Gelgi, M., Guan, Y., Arunachala, S., Samba Siva Rao, M., & Dragoni, N. (2024). Systematic literature review of IoT botnet DDOS attacks and evaluation of detection techniques. Sensors, 24(11), 3571. https://doi.org/10.3390/s24113571

[5] Hurtado Sanchez, J. A., Casilimas, K., & Caicedo Rendon, O. M. (2022). Deep reinforcement learning for resource management on network slicing: A survey. *Sensors*, *22*(8), 3031. https://doi.org/10.3390/s22083031

[6] Zhou, Q., Zhu, J., Zhang, J., Jia, Z., Huberman, B., & Chang, G. K. (2020, May). Intelligent bandwidth allocation for latency management in NG-EPON using reinforcement learning method. In *2020 Conference on Lasers and Electro-Optics (CLEO)* (pp. 1-2). IEEE. http://dx.doi.org/10.48550/arXiv.2001.07698.

[7] Ganesan, E., Hwang, I. S., Liem, A. T., & Ab-Rahman, M. S. (2021, April). 5G-enabled tactile internet resource provision via software-defined optical access networks (SDOANs). In *Photonics* (Vol. 8, No. 5, p. 140). MDPI. https://doi.org/10.3390/photonics8050140.

[8] Arshad, Q. U. A., Khan, W. Z., Azam, F., Khan, M. K., Yu, H., & Zikria, Y. B. (2023). Blockchain-based decentralized trust management in IoT: systems, requirements and challenges. Complex & Intelligent Systems, 9(6), 6155-6176. http://dx.doi.org/10.1007/s40747-023-01058-8

[9] Obaidat, M. A., Rawashdeh, M., Alja'afreh, M., Abouali, M., Thakur, K., & Karime, A. (2024). Exploring IoT and blockchain: a comprehensive survey on security, integration strategies, applications and future research directions. Big Data and Cognitive Computing, 8(12), 174. https://doi.org/10.3390/bdcc8120174

[10] Haque, E. U., Abbasi, W., Almogren, A., Choi, J., Altameem, A., Rehman, A. U., & Hamam, H. (2024). Performance enhancement in blockchain based IoT data sharing using lightweight consensus algorithm. Scientific Reports, 14(1), 26561. https://doi.org/10.1038/s41598-024-77706-x

[11] Wen, D., Bennis, M., & Huang, K. (2020). Joint Parameter-and-Bandwidth Allocation for Improving the Efficiency of Partitioned Edge Learning. *IEEE Transactions on Wireless Communications*, 19, 8272-8286. https://doi.org/10.1109/TWC.2020.3021177.

[12] Hatem, J., Dhaini, A., & Elbassuoni, S. (2019). Deep Learning-Based Dynamic Bandwidth Allocation for Future Optical Access Networks. *IEEE Access*, 7, 97307-97318. https://doi.org/10.1109/ACCESS.2019.2929480.

[13] Wong, E., & Ruan, L. (2023). Towards 6G: fast and self-adaptive dynamic bandwidth allocation for next-generation mobile fronthaul. *Journal of Optical Communications and Networking*, 15, C203-C211. https://doi.org/10.1364/JOCN.483983.

[14] Hu, Y., Li, Q., Chai, Y., Wu, D., Lu, L., Shi, N., Teng, Y., & Zhang, Y. (2024). AI Service Deployment and Resource Allocation Optimization Based on Human-Like Networking Architecture. *IEEE Internet of Things Journal*, 11, 24795-24813. https://doi.org/10.1109/JIOT.2024.3384546.

[15] Liem, A., Hwang, I., Ganesan, E., Taju, S., & Sandag, G. (2023). A Novel Temporal Dynamic Wavelength Bandwidth Allocation Based on Long-Short-Term-Memory in NG-EPON. *IEEE Access*, 11, 82095-82107. https://doi.org/10.1109/ACCESS.2023.3299037.

[16] Żotkiewicz, M., Szałyga, W., Domaszewicz, J., Bąk, A., Kopertowski, Z., & Kozdrowski, S. (2021). Artificial Intelligence Control Logic in Next-Generation Programmable Networks. *Applied Sciences*. https://doi.org/10.3390/app11199163.

[17] Chen, Y. (2023). An adaptive heuristic algorithm to solve the network slicing resource management problem. *International Journal of Communication Systems*, 36. https://doi.org/10.1002/dac.5463.

[18] Chen, W., & Chen, L. (2024). 3.53-TOPS/W EEAIP: An Energy-Efficient Artificial Intelligence Hardware Architecture for Edge AI Applications. *IEEE Transactions on Consumer Electronics*, 70, 4333-4344. https://doi.org/10.1109/TCE.2023.3323644.

[19] Lin, Z., Bi, S., & Zhang, Y. (2020). Optimizing AI Service Placement and Resource Allocation in Mobile Edge Intelligence Systems. *IEEE Transactions on Wireless Communications*, 20, 7257-7271. https://doi.org/10.1109/TWC.2021.3081991.

[20] Qureshi, M., & Tekin, C. (2020). Fast Learning for Dynamic Resource Allocation in AI-Enabled Radio Networks. *IEEE Transactions on Cognitive Communications and Networking*, 6, 95-110. https://doi.org/10.1109/TCCN.2019.2953607.

[21] Iyer, V., Lee, S., Lee, S., Kim, J., Kim, H., & Shin, Y. (2023). Automated Backend Allocation for Multi-Model, On-Device AI Inference. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 7, 1 - 33. https://doi.org/10.1145/3626793.

[22] Goswami, P., Mukherjee, A., Maiti, M., Tyagi, S., & Yang, L. (2021). A Neural-Network-Based Optimal Resource Allocation Method for Secure IIoT Network. *IEEE Internet of Things Journal*, 9, 2538-2544. https://doi.org/10.1109/JIOT.2021.3084636.

[23] Moudoud, H., & Cherkaoui, S. (2023). Empowering Security and Trust in 5G and Beyond: A Deep Reinforcement Learning Approach. *IEEE Open Journal of the Communications Society*, 4, 2410-2420. https://doi.org/10.1109/OJCOMS.2023.3313352.

[24] Bhardwaj, V., Shareef, A., Singh, R., Gharban, H., & Essa, I. (2025). Improving IoT Service Quality with the Allocation of Dynamic Bandwidth. *2025 3rd International Conference on Disruptive Technologies (ICDT)*, 1276-1281. https://doi.org/10.1109/ICDT63985.2025.10986304.

[25] Liang, W., Zhao, J., Liu, Y., Liang, Y., & Li, J. (2023). Fairness resource allocation based on blockchain for secure communication in integrated IoT. *EURASIP Journal on Advances in Signal Processing*, 2023, 1-23. https://doi.org/10.1186/s13634-023-01075-2.

[26] Tsai, W., Wang, S., Liang, Y., & Yang, D. (2022). Optimized Bandwidth Allocation for MEC Server in Blockchain-Enabled IoT Networks. *Scientific Programming*. https://doi.org/10.1155/2022/6129150.

[27] Liao, C., Shuai, H., & Wang, L. (2019). RNN-Assisted Network Coding for Secure Heterogeneous Internet of Things With Unreliable Storage. *IEEE Internet of Things Journal*, 6, 7608-7622. https://doi.org/10.1109/JIOT.2019.2902376.

[28] Abedin, S., Alam, M., Kazmi, S., Tran, N., Niyato, D., & Hong, C. (2019). Resource Allocation for Ultra-Reliable and Enhanced Mobile Broadband IoT Applications in Fog Network. *IEEE Transactions on Communications*, 67, 489-502. https://doi.org/10.1109/TCOMM.2018.2870888.

[29] Rouhifar, M., Hedayati, A., & Aghazarian, V. (2024). DITRA: an efficient event-driven multi-objective optimization algorithm for bandwidth allocation in IoT environments. *Cluster Computing*, 1-21. https://doi.org/10.1007/s10586-023-04214-4.

[30] Chakour, I., Daoui, C., Baslam, M., Sainz-De-Abajo, B., & Garcia-Zapirain, B. (2024). Strategic Bandwidth Allocation for QoS in IoT Gateway: Predicting Future Needs Based on IoT Device Habits. *IEEE Access*, 12, 6590-6603. https://doi.org/10.1109/ACCESS.2024.3351111.

[31] Guan, X., Yu, B., Yu, T., & Cai, Y. (2023). Minimizing age of information in the uplink multi-user networks via dynamic bandwidth allocation. *China Communications*, 20, 287-301. https://doi.org/10.23919/JCC.fa.2022-0586.202304.

[32] Thangavelu, A., & Rajendran, P. (2024). Energy-Efficient Secure Routing for a Sustainable Heterogeneous IoT Network Management. *Sustainability*. https://doi.org/10.3390/su16114756.

[33] Yang, H., Alphones, A., Zhong, W., Chen, C., & Xie, X. (2020). Learning-Based Energy-Efficient Resource Management by Heterogeneous RF/VLC for Ultra-Reliable Low-Latency Industrial IoT Networks. *IEEE Transactions on Industrial Informatics*, 16, 5565-5576. https://doi.org/10.1109/TII.2019.2933867.

[34] Gan, Z., Chen, Y., Xiao, Y., Zhou, D., Feng, C., & Shen, B. (2025). Privacy preservation-driven communication-computing collaboration for multi-mode heterogeneous IoT network management. *IET Commun.*, 19. https://doi.org/10.1049/cmu2.70003.

[35] Sood, K., Karmakar, K., Yu, S., Varadharajan, V., Pokhrel, S., & Xiang, Y. (2020). Alleviating Heterogeneity in SDN-IoT Networks to Maintain QoS and Enhance Security. *IEEE Internet of Things Journal*, 7, 5964-5975. https://doi.org/10.1109/JIOT.2019.2959025.

[36] Selvaraju, S., Balador, A., Fotouhi, H., Vahabi, M., & Björkman, M. (2021). Network Management in Heterogeneous IoT Networks. *2021 International Wireless Communications and Mobile Computing (IWCMC)*, 1581-1586. https://doi.org/10.1109/IWCMC51323.2021.9498801.

[37] Tseng, L., Wong, L., Otoum, S., Aloqaily, M., & Ben-Othman, J. (2020). Blockchain for Managing Heterogeneous Internet of Things: A Perspective Architecture. *IEEE Network*, 34, 16-23. https://doi.org/10.1109/MNET.001.1900103.

[38] Kherraf, N., Alameddine, H., Sharafeddine, S., Assi, C., & Ghrayeb, A. (2019). Optimized Provisioning of Edge Computing Resources With Heterogeneous Workload in IoT Networks. *IEEE Transactions on Network and Service Management*, 16, 459-474. https://doi.org/10.1109/TNSM.2019.2894955.

[39] Khan, A., Rizwan, A., Ahmad, R., Jin, W., Khan, Q., Lim, S., & Kim, D. (2024). Hetero-FedIoT: A Rule-Based Interworking Architecture for Heterogeneous Federated IoT Networks. *IEEE Internet of Things Journal*, 11, 5920-5938. https://doi.org/10.1109/JIOT.2023.3308579.

[40] Zafar, A., Samad, F., Syed, H., Ibrahim, A., Alohaly, M., & Elsadig, M. (2023). An Advanced Strategy for Addressing Heterogeneity in SDN-IoT Networks for Ensuring QoS. *Applied Sciences*. https://doi.org/10.3390/app13137856.

[41] Hussain, F., Hassan, S., Hussain, R., & Hossain, E. (2019). Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges. *IEEE Communications Surveys & Tutorials*, 22, 1251-1275. https://doi.org/10.1109/COMST.2020.2964534.

[42] Javadpour, A., Wang, G., & Rezaei, S. (2020). Resource Management in a Peer to Peer Cloud Network for IoT. *Wireless Personal Communications*, 115, 2471 - 2488. https://doi.org/10.1007/s11277-020-07691-7.

[43] Kadiyala, R., Narayana, C. L., Ramu, S. C., Putta, N., Pabboju, S. S., & Reddy, B. R. (2024). Trust-Aware Federated Learning with Differential Privacy for Secure AIoT in Critical Infrastructures. *EAI Endorsed Transactions on Internet of Things*, *11*.