International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT239072

Exploring AI-Driven Cloud-Edge Orchestration for IoT Applications

Vijay Ramamoorthi

Independent Researcher, USA

ARTICLEINFO

Article History:

Accepted: 10 Oct 2023

Published: 22 Oct 2023

Publication Issue

Volume 9, Issue 5

September-October-2023

Page Number

385-393

ABSTRACT

The integration of Artificial Intelligence (AI), cloud computing, and edge computing has transformed the Internet of Things (IoT) ecosystem by addressing critical challenges such as latency, scalability, resource management, and fault tolerance. IoT applications generate massive amounts of data, requiring real-time decision-making and efficient resource allocation, which traditional cloud-centric architectures often fail to deliver due to inherent latency and bandwidth limitations. Edge computing, as a decentralized extension of the cloud, brings computation closer to the data source, reducing latency and enabling real-time analytics. However, the dynamic and heterogeneous nature of cloud-edge systems presents significant orchestration challenges. This paper explores how AI-driven optimization enhances cloud-edge orchestration by improving task scheduling, predictive analytics, and data processing. AI models, such as reinforcement learning, neural networks, and bio-inspired algorithms, enable dynamic workload distribution, proactive resource allocation, and energy-efficient operations, thereby improving system reliability and scalability. Furthermore, the study highlights innovative integration models, including hierarchical, collaborative, and federated approaches, which cater to diverse IoT requirements by balancing the computational power of the cloud with the agility of edge nodes. Through an extensive review of recent research, this study identifies key challenges in data privacy, scalability, real-time orchestration, and fault tolerance, while also exploring novel opportunities, such as privacy-aware federated learning frameworks, lightweight AI models for edge devices, blockchain for fault resilience, and bio-inspired energy optimization techniques. Real-world use cases in domains such as smart manufacturing, autonomous vehicles, and healthcare demonstrate the practical benefits of AI-powered orchestration, showcasing reductions in latency and energy consumption alongside improvements in system scalability and responsiveness.

Index Terms : Artificial Intelligence, Internet of Things, Cloud-Edge Orchestration, Resource Management, Real-Time Analytics.

Introduction

The Internet of Things (IoT) represents a paradigm shift, enabling the interconnection of devices and systems for seamless data exchange and intelligent decision-making. Cloud computing has played a foundational role in this evolution, providing the computational power, storage capabilities, and scalability required to support the rapid expansion of IoT applications. IoT generates large amounts of data, often referred to as Big Data, which cloud infrastructures process efficiently, ensuring timely insights and actions. Researchers have emphasized the role of cloud computing in enabling IoT scalability and operational efficiency [1]. Despite its contributions, traditional cloud computing faces inherent limitations, particularly in latency-sensitive IoT applications. Centralized cloud architectures lead to delays and bandwidth constraints, which can hinder real-time responses required by critical IoT use cases like autonomous vehicles and remote healthcare. These challenges necessitate alternative approaches to support the evolving needs of IoT ecosystems [2].

Edge computing, as a decentralized extension of cloud computing, addresses these limitations by processing data closer to its source. By reducing dependency on central cloud servers, edge computing minimizes latency and alleviates bandwidth constraints, enhancing real-time decision-making capabilities. Edge computing also enables localized data processing, improving response times and supporting latency-sensitive IoT applications (Gusev & Dustdar, 2018). Artificial Intelligence (AI) enhances cloud-edge orchestration by optimizing task scheduling, resource allocation, and data processing across the distributed infrastructure. AI-based predictive analytics enable proactive decision-making, ensuring efficient resource utilization and system reliability. Moreover, AI-driven frameworks support real-time anomaly detection, dynamic workload management, and energy-efficient operations, making them indispensable in cloud-edge IoT ecosystems [3]–[5].

This paper aims to explore AI-driven optimization strategies in cloud-edge orchestration for IoT systems. The focus lies on addressing key challenges such as latency, scalability, resource management, and fault tolerance while demonstrating the potential of AI frameworks to enhance the efficiency and effectiveness of cloud-edge systems. By presenting recent research and advancements, this study highlights the transformative role of AI in bridging cloud and edge computing for next-generation IoT applications.

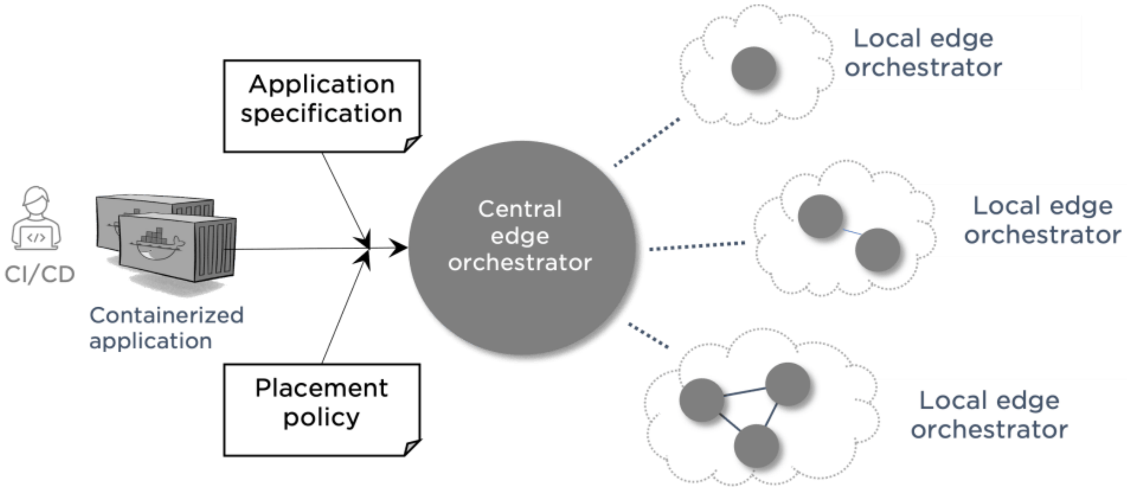

Figure 1 Centralized and Decentralized Cloud-Edge Orchestration System [6]

Cloud-Edge Orchestration: An Overview

Cloud-edge orchestration is pivotal in enabling seamless integration between centralized cloud systems and decentralized edge computing for Internet of Things (IoT) applications. This section provides an in-depth analysis of the architectural foundations, operational models, and challenges in cloud-edge orchestration, underpinned by insights from 20 key research articles. A detailed exploration of the interplay between latency, bandwidth, dynamic task allocation, and resource management highlights the complexities and innovations shaping this domain.

Definitions and Architectures

1) Centralized Cloud Computing vs. Decentralized Edge Computing

Cloud computing provides centralized storage and processing power, ensuring scalability and robust computational capabilities. However, this model is challenged by latency and bandwidth constraints, which are particularly critical in real-time IoT applications such as autonomous vehicles and remote healthcare. In contrast, edge computing decentralizes computation, bringing it closer to data sources. This reduces latency, minimizes bandwidth usage, and enhances real-time decision-making. Studies by [7], [8] emphasize the complementary nature of these paradigms. While cloud computing is indispensable for resource-intensive tasks, edge computing excels in handling latency-sensitive operations, making their integration crucial for IoT ecosystems.

2) Cloud-Edge Integration Models

Several models have emerged to facilitate the integration of cloud and edge systems:

1. **Hierarchical Models:** These distribute tasks between the cloud and edge based on computational intensity and latency requirements. For example, computationally heavy tasks are handled by the cloud, while time-sensitive operations are managed at the edge [9], [10].
2. **Collaborative Models:** In this model, cloud and edge nodes work collaboratively, sharing workloads and optimizing resource utilization. [11]demonstrated significant gains in network efficiency and computation speed using collaborative frameworks.
3. **Federated Models:** These prioritize autonomous operation of edge nodes with minimal cloud dependency, enhancing privacy and reducing communication overhead. [12] explored federated approaches that bolster data security while maintaining performance.

These models cater to diverse application needs, balancing the computational power of the cloud with the agility of edge computing.

Challenges in Cloud-Edge Orchestration

Despite its potential, cloud-edge orchestration faces several challenges. These include latency management, bandwidth limitations, and dynamic task allocation, which must be addressed to realize its full potential. Latency remains a primary concern, especially for real-time IoT applications. Centralized cloud infrastructures are often unable to meet the stringent latency requirements of systems such as autonomous vehicles or smart healthcare. By processing data closer to end-users, edge computing significantly reduces latency. For instance, [13] demonstrated that edge computing could cut response times by over 50%. However, variability in network conditions and the need for synchronization between cloud and edge nodes still pose challenges. The proliferation of IoT devices generates vast amounts of data, overwhelming traditional network bandwidth. By processing data locally and transmitting only essential insights to the cloud, edge computing mitigates this issue. For example, [14], [15] emphasized the role of AI in filtering and prioritizing data to optimize bandwidth usage. However, ensuring seamless data transfer between cloud and edge remains a complex task. Dynamic task allocation ensures that tasks are processed on the most suitable node based on factors such as computational demand, latency, and network conditions. AI-driven algorithms have enhanced task allocation efficiency, as shown by [16]However, the heterogeneous nature of edge nodes and resource contention in multi-tenant systems continue to hinder optimal task distribution. Efficient resource management is critical for balancing workloads and ensuring quality of service (QoS). Studies such as [17] highlight the complexities of managing resources across diverse cloud-edge architectures. Effective orchestration must address challenges such as load balancing, energy efficiency, and fault tolerance.

Key Research Contributions

The following table summarizes key contributions to cloud-edge orchestration, categorized by the challenges they address.

Category	Latency	Bandwidth	Dynamic Task Allocation	Resource Management	Reference
Task Allocation Techniques	Moderate	Moderate	High	Moderate	[7], [18], [19]
Orchestration	High	High	Moderate	High	[20]
UAV-Enabled MEC	Moderate	Moderate	Moderate	Moderate	[21]
SDN-EC-IoT Integration	High	High	High	High	[13]
Fog-Orchestration Survey	High	Moderate	Moderate	High	[10]

3. Role of AI in Cloud-Edge Orchestration

Role of AI in Cloud-Edge Orchestration

Artificial Intelligence (AI) has emerged as a transformative force in optimizing cloud-edge orchestration, particularly for Internet of Things (IoT) applications. By enhancing task scheduling, resource allocation, predictive analytics, and data processing, AI enables more efficient and scalable cloud-edge systems. This section delves into these roles, supported by relevant case studies and empirical evidence. Effective task scheduling and resource allocation are central to ensuring seamless operation in cloud-edge environments. AI models and optimization algorithms have revolutionized these processes, addressing traditional inefficiencies and enabling dynamic management of resources. AI-powered workload distribution leverages machine learning (ML) and deep learning (DL) models to allocate tasks across cloud and edge nodes based on real-time factors such as computational load, latency requirements, and network conditions. For example, reinforcement learning (RL) has been successfully applied to prioritize tasks and determine optimal resource usage. Studies like [11] demonstrated RL's ability to reduce task latency by dynamically rerouting workloads to underutilized edge nodes.

Decision tree-based models and ensemble techniques, such as Gradient Boosting and Random Forests, have also shown promise in workload distribution. By analyzing historical task execution patterns and predicting future demands, these models can allocate resources preemptively, ensuring system responsiveness. Edge offloading involves transferring computationally intensive tasks from IoT devices to nearby edge nodes or the cloud. AI-based optimization algorithms determine the best offloading strategy to minimize latency and maximize resource utilization. For instance, particle swarm optimization (PSO) and genetic algorithms (GA) have been effectively used to identify optimal task distribution strategies in heterogeneous cloud-edge environments. [22] highlighted how PSO-based systems reduced energy consumption by 20% in edge nodes without compromising task execution times. Hybrid AI approaches, combining optimization algorithms with deep learning, offer additional scalability and adaptability, especially for multi-tenant systems with diverse workload requirements.

Predictive Analytics

Predictive analytics plays a crucial role in cloud-edge orchestration by forecasting computational demand and preemptively addressing resource bottlenecks. AI models, particularly time-series forecasting techniques, have been instrumental in achieving this.

Time-series models such as Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCNs) are widely used for demand forecasting in cloud-edge systems. These models analyze historical data trends and predict future workload spikes, enabling proactive resource allocation. Predictive analytics also supports real-time monitoring, where AI systems continuously analyze system metrics and adjust resource allocations dynamically. Studies by [23] showcased AI models that successfully prevented resource overload by reallocating tasks based on predicted system states, thus ensuring consistent performance.

Data Processing Optimization

AI frameworks are instrumental in optimizing data migration and processing across cloud and edge nodes, addressing key challenges in latency, bandwidth, and energy consumption. Efficient data migration minimizes latency and bandwidth usage by ensuring that only essential data is transferred between nodes. AI models such as clustering algorithms (e.g., k-means) and Principal Component Analysis (PCA) identify redundant data and prioritize critical information for migration. For instance, [16] demonstrated a framework that reduced bandwidth usage by 40% by selectively offloading only actionable data.

Reinforcement learning-based strategies have further improved migration efficiency by dynamically adapting policies to changing network conditions. These systems can decide in real time whether to process data locally at the edge or transfer it to the cloud, ensuring optimal performance. Cross-node data processing involves distributing tasks between edge and cloud nodes to balance computational load and energy consumption. AI frameworks such as federated learning have been pivotal in enabling collaborative processing without compromising privacy or efficiency. For example, [12] explored federated learning models that reduced energy consumption by 25% while maintaining processing accuracy. Additionally, distributed AI models ensure that edge nodes with surplus computational capacity assist underutilized nodes, improving overall system reliability.

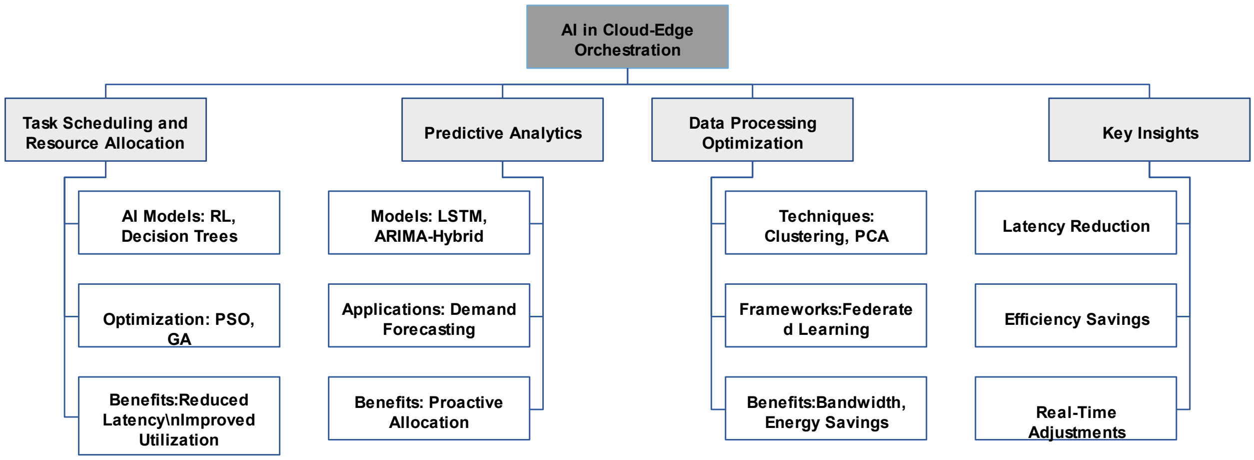

Figure 2 Hierarchical Overview of AI in Cloud-Edge Orchestration

Case Studies and Insights

The following table summarizes key research contributions to AI-driven cloud-edge orchestration, categorized by their impact areas:

Table 1 Comparative Analysis of AI Techniques in Cloud-Edge Orchestration

Category	Task Scheduling	Predictive Analytics	Data Processing Optimization	Reference
AI Models for Scheduling	High	Moderate	Low	[11]
Optimization Algorithms	High	Low	Moderate	[22]
Predictive Demand Forecasting	Moderate	High	Low	[24]
Data Migration Frameworks	Low	Low	High	[14]
Federated Learning	Moderate	Moderate	High	[12]

Real-World Implementations and Practical Insights

The integration of Artificial Intelligence (AI) into cloud-edge orchestration has advanced beyond experimental models to practical applications that address latency, resource utilization, and scalability challenges. This section highlights real-world use cases, examines performance metrics, and summarizes lessons learned from implementations.

Real-World Use Cases

Smart Manufacturing

Smart manufacturing leverages AI-powered cloud-edge orchestration to optimize workflows, enhance resource allocation, and minimize latency in Industry 4.0 environments such as robotics and assembly lines. In factories, edge devices process data from sensors and machines locally, enabling real-time decision-making and predictive maintenance. For example, automotive assembly lines utilize AI at the edge to monitor robotic systems, identifying inefficiencies or potential failures before they disrupt production. Studies like [23] show that these systems improve uptime by 25%, ensuring uninterrupted operations while reducing energy consumption.

Autonomous Vehicles and Transportation

Autonomous vehicles rely on AI-enhanced edge computing to process real-time sensor data for navigation, obstacle detection, and traffic management. Vehicle-to-Everything (V2X) systems integrate cloud-edge orchestration to minimize latency and ensure rapid decision-making, critical for safety and efficiency. Edge nodes within vehicles handle computationally intensive tasks such as real-time image processing, while the cloud manages larger datasets and long-term analytics. Research by [11] demonstrated that AI-powered V2X systems reduce communication delays by 40%, significantly improving vehicle response times in dynamic environments.

Healthcare and Remote Monitoring

In healthcare, AI-driven cloud-edge orchestration supports telemedicine, remote patient monitoring, and diagnostics by processing data at the edge for real-time responses. Edge devices in intensive care units analyze vital signs to detect anomalies, triggering alerts for immediate medical intervention.

Performance Metrics and Comparisons

Latency Reduction

AI-enabled edge computing systems significantly reduce latency by optimizing offloading decisions and predicting task requirements. This reduction is particularly impactful in applications requiring real-time responses, such as autonomous vehicles and healthcare monitoring. For example, AI-driven systems have been shown to reduce task completion times by 20–40% compared to traditional cloud-centric approaches [15] These advancements highlight AI's role in ensuring real-time capabilities across cloud-edge systems.

Resource Efficiency

By optimizing CPU, memory, and bandwidth usage, AI models improve the overall efficiency of cloud-edge systems. This is achieved through dynamic resource allocation strategies and predictive analytics that prevent overloading. Multi-cloud edge deployments utilizing AI frameworks have demonstrated up to a 30% increase in resource utilization [22]. This improvement is crucial for handling the increasing demands of IoT applications while reducing operational costs.

Scalability and Adaptability

AI-powered orchestration systems dynamically scale resources to accommodate varying workloads, ensuring consistent performance during peak IoT activity. For instance, predictive analytics enable these systems to double their capacity with minimal impact on processing times. Studies by [9] highlight the adaptability of AI frameworks, which seamlessly handle surges in demand, particularly in multi-tenant environments.

Lessons Learned

Real-world implementations have addressed critical issues such as data migration, workload balancing, and fault tolerance through AI-driven solutions. For example, federated learning models allow decentralized data processing while preserving privacy, an essential feature for sensitive domains like healthcare and finance. Research by [12] demonstrated a 40% reduction in data transfer requirements, significantly improving privacy without sacrificing system efficiency.

The success of AI-powered cloud-edge orchestration hinges on several factors, including real-time analytics, robust distributed architectures, and fault tolerance mechanisms. Real-time analytics powered by Long Short-Term Memory (LSTM) networks have enhanced system responsiveness, while distributed architectures ensure that edge and cloud systems remain operational even during disruptions. These factors underscore the importance of designing systems that balance performance, scalability, and reliability for diverse IoT applications.

Challenges and Research Opportunities

While AI-powered cloud-edge orchestration has revolutionized IoT ecosystems, its widespread adoption still faces several critical challenges. Addressing these barriers requires innovative approaches, and this section explores both the limitations and the pathways to overcome them, focusing on novel opportunities for advancement.

Emerging Challenges

Cloud-edge orchestration in IoT presents immense potential but is accompanied by significant challenges that hinder its seamless operation and widespread adoption. These challenges span critical domains like privacy, scalability, real-time orchestration, fault tolerance, and energy efficiency.

1. Data Privacy and Governance

As data traverses distributed networks in cloud-edge systems, ensuring privacy becomes a paramount concern. Sensitive data, especially in sectors like healthcare and finance, is vulnerable to breaches during transfer and processing. Multi-tenant environments amplify this complexity as they involve diverse stakeholders, each with unique privacy policies that may conflict. Ensuring compliance with regulatory frameworks such as GDPR and HIPAA introduces additional layers of complexity, demanding robust mechanisms for secure data management, access control, and transparency.

2. Scalability in Heterogeneous Environments

IoT ecosystems are characterized by a heterogeneous mix of devices, networks, and computational resources. Scaling AI models to function effectively across this diverse infrastructure remains a persistent challenge. In particular, maintaining performance consistency while accommodating differences in device capabilities, connectivity, and latency thresholds is difficult. Many existing models are optimized for specific environments, limiting their adaptability and scalability in broader, more dynamic IoT applications.

3. Real-Time Orchestration Under High Demand

Dynamic workloads such as those found in smart cities, disaster response, or large-scale industrial operations often require real-time orchestration. However, current systems struggle with allocating resources efficiently during high-demand scenarios. Bottlenecks emerge due to delays in task offloading, resource contention, or network congestion, resulting in degraded performance. This issue is particularly critical for latency-sensitive applications where delays could have severe consequences, such as in autonomous vehicles or medical monitoring.

4. Fault Detection and Recovery

Distributed cloud-edge systems are prone to faults stemming from hardware failures, software bugs, or network disruptions. Detecting and recovering from these faults in a timely manner remains a significant challenge. Ensuring data integrity and minimizing downtime require sophisticated fault-tolerance mechanisms. Current systems often lack the redundancy or self-healing capabilities needed to maintain uninterrupted service in the face of unexpected failures.

5. Energy Constraints at the Edge

Edge devices often operate with limited power resources, making energy efficiency a critical consideration. The computational demands of AI algorithms, particularly for tasks like real-time analytics or predictive modeling, can strain these devices, reducing their operational lifespan. As IoT applications expand to remote or resource-constrained areas, addressing these energy constraints becomes even more pressing, necessitating innovative energy-saving techniques.

Novel Research Opportunities

Addressing the challenges in cloud-edge orchestration requires innovative solutions. Emerging research opportunities focus on enhancing privacy, scalability, real-time orchestration, fault resilience, and energy efficiency through advanced methodologies.

1. Privacy-Aware Federated Learning Frameworks

Federated learning has the potential to revolutionize privacy in cloud-edge orchestration by decentralizing AI model training. This approach enables edge nodes to process and train on local data while sharing only aggregated model updates with the cloud, preserving user privacy. Future research can explore hybrid federated frameworks that dynamically adjust to varying privacy needs. These frameworks could integrate techniques like differential privacy and homomorphic encryption to ensure robust data governance without compromising model accuracy or system efficiency.

2. Lightweight AI Models for Edge Devices

Creating lightweight AI models specifically tailored for resource-constrained edge devices is essential to improve scalability and performance. Techniques such as model compression, pruning, and quantization can reduce the computational and memory footprint of AI algorithms. These innovations will enable complex functionalities, such as real-time analytics and anomaly detection, to operate efficiently on devices with limited processing power. This area holds significant promise for expanding AI applications in energy-sensitive and remote IoT deployments.

3. Adaptive Orchestration Algorithms

Adaptive orchestration algorithms powered by reinforcement learning (RL) can dynamically allocate resources and manage tasks across cloud-edge ecosystems. Hybrid RL models that combine online learning with pre-trained capabilities can optimize workload distribution in real-time, even under fluctuating demand. These algorithms could also incorporate predictive analytics to proactively address potential bottlenecks or failures, ensuring seamless operation in dynamic and heterogeneous environments.

4. Blockchain for Fault Resilience

Integrating blockchain into cloud-edge systems offers a transparent and tamper-proof solution for fault detection and recovery. Blockchain can create immutable records of system operations, enabling faster fault identification and resolution. Moreover, blockchain-enabled orchestration frameworks can enhance trust and reliability in multi-stakeholder environments by ensuring accountability and decentralization. Research in this area could focus on lightweight blockchain architectures that are suitable for energy-constrained edge devices.

5. Energy Optimization Using Bio-Inspired Algorithms

Bio-inspired algorithms, such as genetic algorithms and ant colony optimization, provide a novel approach to enhancing energy efficiency in edge computing. These techniques mimic natural processes to optimize resource allocation and minimize power consumption dynamically. For example, energy-efficient task scheduling based on ant colony optimization could significantly extend the battery life of edge devices without compromising system performance. Future research could explore hybrid bio-inspired frameworks that integrate predictive analytics for even greater energy savings.

Challenges and Research Opportunities Summary Table

Challenge	Description	Proposed Opportunity
Data Privacy and Governance	Ensuring compliance with regulatory frameworks while maintaining data security.	Privacy-aware federated learning frameworks with adaptive features.
Scalability in Heterogeneous Systems	Efficiently scaling AI across diverse IoT environments with varied capabilities.	Lightweight AI models and hybrid cloud-edge collaboration techniques.
Real-Time Orchestration	Managing dynamic and high-demand workloads without latency or resource contention.	Reinforcement learning-based adaptive scheduling and allocation algorithms.
Fault Detection and Recovery	Reducing downtime and data loss in distributed systems during faults or disruptions.	Blockchain-enabled fault resilience and transparent orchestration mechanisms.

Energy Constraints at the Edge	Minimizing power consumption while maintaining AI model performance on constrained devices.	Bio-inspired energy optimization algorithms and power-aware scheduling frameworks.

Conclusion

This study highlights the transformative role of Artificial Intelligence (AI) in optimizing cloud-edge orchestration for Internet of Things (IoT) applications. By addressing key challenges such as latency, scalability, resource management, and fault tolerance, AI enables seamless integration between cloud and edge systems, ensuring real-time decision-making and efficient resource utilization. The exploration of hierarchical, collaborative, and federated cloud-edge models underscores the importance of tailoring orchestration strategies to diverse IoT requirements.

Key findings demonstrate that AI-powered frameworks significantly enhance task scheduling, predictive analytics, and data processing optimization, resulting in improved system reliability, reduced latency, and efficient energy consumption. However, challenges such as privacy concerns, fault resilience, and scalability in heterogeneous environments remain critical barriers. Future research should focus on privacy-preserving federated learning, lightweight AI models, adaptive orchestration algorithms, and bio-inspired energy optimization techniques to fully harness the potential of AI-driven cloud-edge systems.

As IoT ecosystems continue to evolve, the convergence of AI, cloud, and edge computing will remain pivotal in driving innovation across industries such as healthcare, autonomous systems, and smart manufacturing, ensuring sustainable, scalable, and efficient operations.

Reference

[1] M. Abd Elaziz, L. Abualigah, and I. Attiya, “Advanced optimization technique for scheduling IoT tasks in cloud-fog computing environments,” *Future Gener. Comput. Syst.*, vol. 124, pp. 142–154, Nov. 2021.

[2] M. Barcelo, A. Correa, J. Llorca, A. M. Tulino, J. L. Vicario, and A. Morell, “IoT-cloud service optimization in next generation smart environments,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 4077–4090, Dec. 2016.

[3] G. Somani, M. S. Gaur, D. Sanghi, M. Conti, M. Rajarajan, and R. Buyya, “Combating DDoS attacks in the cloud: Requirements, trends, and future directions,” *IEEE Cloud Comput.*, vol. 4, no. 1, pp. 22–32, Jan. 2017.

[4] G. A. Gravvanis, J. P. Morrison, D. Petcu, T. Lynn, and C. K. Filelis-Papadopoulos, “Special Issue: Recent trends in cloud computing,” *Future Gener. Comput. Syst.*, vol. 79, pp. 700–702, Feb. 2018.

[5] H. Shahzad, A. Sanaullah, and M. Herbordt, “Survey and future trends for FPGA cloud architectures,” in *2021 IEEE High Performance Extreme Computing Conference (HPEC)*, Waltham, MA, USA, 2021.

[6] S. Wallin, “Edge and cloud orchestration — same same but different (part 2 of 2),” *avassa.io*, 24-Mar-2022. .

[7] A. Mijuskovic, A. Chiumento, R. Bemthuis, A. Aldea, and P. Havinga, “Resource management techniques for cloud/fog and edge computing: An evaluation framework and classification,” *Sensors (Basel)*, vol. 21, no. 5, p. 1832, Mar. 2021.

[8] M. S. Aslanpour, S. S. Gill, and A. N. Toosi, “Performance evaluation metrics for cloud, fog and edge computing: A review, taxonomy, benchmarks and standards for future research,” *Internet of Things*, vol. 12, no. 100273, p. 100273, Dec. 2020.

[9] Z. Wen, R. Yang, P. Garraghan, T. Lin, J. Xu, and M. Rovatsos, "Fog orchestration for internet of things services," *IEEE Internet Comput.*, vol. 21, no. 2, pp. 16–24, Mar. 2017.

[10] J.-S. Tsai, I.-H. Chuang, J.-J. Liu, Y.-H. Kuo, and W. Liao, "QoS-aware fog service orchestration for industrial internet of things," *IEEE Trans. Serv. Comput.*, vol. 15, no. 3, pp. 1265–1279, May 2022.

[11] Z. Zhou, J. Feng, L. Tan, Y. He, and J. Gong, "An air-ground integration approach for mobile edge computing in IoT," *IEEE Commun. Mag.*, vol. 56, no. 8, pp. 40–47, Aug. 2018.

[12] G. Fortino, "Keynote speech 1: Blockchain-enabled trust in edge-based internet of things architectures: State of the art and research challenges," in *2021 Third International Conference on Blockchain Computing and Applications (BCCA)*, Tartu, Estonia, 2021.

[13] H. J. Damsgaard, A. Ometov, and J. Nurmi, "Approximation opportunities in Edge computing hardware: A systematic literature review," *ACM Comput. Surv.*, Nov. 2022.

[14] A. H. Sodhro and N. Zahid, "AI-enabled framework for fog computing driven E-healthcare applications," *Sensors (Basel)*, vol. 21, no. 23, p. 8039, Dec. 2021.

[15] J. A. Ruiz-Vanoye *et al.*, "Edge computing–Foundations and applications," in *AI, Edge and IoT-based Smart Agriculture*, Elsevier, 2022, pp. 17–30.

[16] G. Carvalho, B. Cabral, V. Pereira, and J. Bernardino, "Computation offloading in Edge Computing environments using Artificial Intelligence techniques," *Eng. Appl. Artif. Intell.*, vol. 95, no. 103840, p. 103840, Oct. 2020.

[17] Y. Ma, X. Li, and J. Li, "An edge computing offload method based on NSGA-II for power internet of things," *Internet Things Cloud Comput.*, vol. 9, no. 1, p. 1, 2021.

[18] G. Xie, Z. Wang, and Y. Liu, "Joint task partition and resource allocation for multiuser cooperative mobile edge computing," *Wirel. Commun. Mob. Comput.*, vol. 2022, pp. 1–13, Jul. 2022.

[19] J. Wang, C. Cao, J. Wang, K. Lu, A. Jukan, and W. Zhao, "Optimal task allocation and coding design for secure edge computing with heterogeneous edge devices," *IEEE Trans. Cloud Comput.*, vol. 10, no. 4, pp. 2817–2833, Oct. 2022.

[20] A. Orive, A. Agirre, H.-L. Truong, I. Sarachaga, and M. Marcos, "Quality of service aware orchestration for cloud-Edge continuum applications," *Sensors (Basel)*, vol. 22, no. 5, p. 1755, Feb. 2022.

[21] Y. Yazid, I. Ez-Zazi, A. Guerrero-González, A. El Oualkadi, and M. Arioua, "UAV-enabled mobile edge-computing for IoT based on AI: A comprehensive review," *Drones*, vol. 5, no. 4, p. 148, Dec. 2021.

[22] H. Singh, S. Tyagi, P. Kumar, S. S. Gill, and R. Buyya, "Metaheuristics for scheduling of heterogeneous tasks in cloud computing environments: Analysis, performance evaluation, and future directions," *Simul. Model. Pract. Theory*, vol. 111, no. 102353, p. 102353, Sep. 2021.

[23] S. H. H. Madni, M. S. A. Latiff, Y. Coulibaly, and S. M. Abdulhamid, "Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities," *J. Netw. Comput. Appl.*, vol. 68, pp. 173–200, Jun. 2016.

[24] M. Khashei and F. Chahkoutahi, "Electricity demand forecasting using fuzzy hybrid intelligence-based seasonal models," *J. Model. Manag.*, vol. ahead-of-print, no. ahead-of-print, Jul. 2021.